

2006

Differential item and test functioning of the Brief Child and Family Phone Interview in First Nations and non-First Nations children and adolescents / by Dana Bova.

Bova, Dana Louise

<http://knowledgecommons.lakeheadu.ca/handle/2453/3326>

Downloaded from Lakehead University, Knowledge Commons

Running head: DIFFERENTIAL FUNCTIONING AND BCFPI

Differential Item and Test Functioning of the Brief Child and Family Phone
Interview in First Nations and non-First Nations Children and Adolescents

Dana L. Bova

Ph. D. Dissertation

Department of Psychology

Lakehead University

March 2006



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-21544-9

Our file Notre référence

ISBN: 978-0-494-21544-9

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Acknowledgements

This research endeavour would not have been possible without the following people. I would like to thank Dr. Brian O'Connor, my dissertation supervisor, for his guidance and wisdom during this research project. I would also like to thank Dr. Dwight Mazmanian for his expertise in assessment issues; Dr. Charles Netley and Dr. Keith Brownlee for their clinical expertise in working with First Nations children and families; and Dr. Darcy Santor for his expertise in item response theory. I am thankful for Dilico Ojibway Child and Family Services agreeing to participate in this research project. I am grateful for the assistance of Dr. Michael Boyle and Yvonne Racine for providing the raw data from the Revised Ontario Child Health Study. Many, many thanks are extended to my family, peeps, and four-legged peeps for their encouragement and continual support. Lastly, I would like to thank Wesley for his consistent love, assistance, guidance, emotional support, and understanding the importance of this personal endeavour.

This research project was supported by the University of Toronto/McMaster University Indigenous Health Research Development Program Graduate Award, funded by the Canadian Institutes of Health Research-Institute of Aboriginal People's Health.

Table of Contents

Abstract	7
Introduction	8
Canada's First Nations People	9
Demographics of First Nations People	9
Historical and Contemporary Issues of First Nations People	10
Mental Health and First Nations People	12
Culture and Clinical Assessment	18
Brief Child and Family Phone Interview	21
Item Response Theory	28
Fundamentals of Item Response Theory	28
Item Response Theory and Item Characteristic Curves	31
Differential Functioning	34
Differential item functioning	35
Differential test functioning	38
Differential functioning analyses	40
Advantages of Item Response Theory	43
Disadvantages of Item Response Theory	44
Purpose of the Present Study	45
Method	47
Participants	47
Measure	48
Results	53
Sample Characteristics	53
Group Differences In Raw Score Means	54
Item Response Theory Nonparametric Analyses	57
Item Response Theory Parametric Analyses	58
Discussion	64
References	80

List of Tables

1. BCFPI Mental Health Scales and Items.	89
2. Demographic Information for First Nations and non-First Nations Participants.	91
3. BCFPI Scale Scores, t Values, and Effect Size for the First Nations Clinical Sample and non-First Nations Clinical Sample.	94
4. BCFPI Scale Scores, t Values, and Effect Size for the First Nations Clinical Sample and non-First Nations General Population Sample.	95
5. BCFPI Scale Scores, t Values, and Effect Size for the non-First Nations Clinical Sample and non-First Nations General Population Sample.	96
6. Differential Item Functioning Among the First Nations Clinical Sample and the non-First Nations Clinical Sample.	97
7. Differential Item Functioning Among the First Nations Clinical Sample and the non-First Nations General Population Sample.	99
8. Differential Test Functioning Among the First Nations Clinical Sample and the non-First Nations Clinical Sample.	101
9. Differential Test Functioning Among the First Nations Clinical Sample and the non-First Nations General Population Sample.	102

List of Figures

1. Flat, Upward, and Downward Slopes.	103
2. Example of hypothetical item characteristic curves between First Nations and non-First Nations children and adolescents on the "Has difficulty following directions or instructions" item from the Regulating Attention, Impulsivity, and Activity Level scale of the BCFPI.	104
3. Example of hypothetical uniform differential item functioning between First Nations and non-First Nations children and adolescents on the "Has difficulty following directions or instructions" item from the Regulating Attention, Impulsivity, and Activity Level scale of the BCFPI.	105
4. Example of hypothetical non-uniform differential item functioning between First Nations and non-First Nations children and adolescents on the "Has difficulty following directions or instructions" item from the Regulating Attention, Impulsivity, and Activity Level scale of the BCFPI.	106
5. An illustration of test characteristic curves for no differential test functioning on the Regulation of Attention, Impulsivity, and Activity Level scale from the BCFPI between First Nations and non-First Nations children and adolescents.	107
6. An illustration of test characteristic curves for differential test functioning on the Regulation of Attention, Impulsivity, and Activity Level scale from the BCFPI between First Nations and non-First Nations children and adolescents.	108
7. Item characteristic curves for each item on the Regulation of Attention, Impulsivity, and Activity scale on the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample.	109
8. Item characteristic curves for each item on the Cooperativeness with Others scale on the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample.	110
9. Item characteristic curves for each item on the Conduct scale on the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample.	111

10. Item characteristic curves for each item on the Separation from Parents scale on the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample.	112
11. Item characteristic curves for each item on the Managing Anxiety scale on the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample.	113
12. Item characteristic curves for each item on the Managing Mood scale on the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample.	114
13. Item characteristic curves for each item on the Regulating Attention, Impulsivity, and Activity scale on the BCFPI among the First Nations clinical sample and the non-First Nations community sample.	115
14. Item characteristic curves for each item on the Cooperativeness with Others scale on the BCFPI among the First Nations clinical sample and the non-First Nations community sample.	116
15. Item characteristic curves for each item on the Conduct scale on the BCFPI among the First Nations clinical sample and the non-First Nations community sample.	117
16. Item characteristic curves for each item on the Separation from Parents scale on the BCFPI among the First Nations clinical sample and the non-First Nations community sample.	118
17. Item characteristic curves for each item on the Managing Anxiety scale on the BCFPI among the First Nations clinical sample and the non-First Nations community sample.	119
18. Item characteristic curves for each item on the Managing Mood scale on the BCFPI among the First Nations clinical sample and the non-First Nations community sample.	120

Abstract

The purpose of the present study was to utilize item response theory techniques to examine bias on Brief Child and Family Phone Interview mental health scales among First Nations and non-First Nations children and adolescents. The participants consisted of 3,265 children and adolescents. There were 340 First Nations and 2925 non-First Nations participants. The First Nations children and adolescents were selected from Dilico Ojibway Child and Family Services, a mental health agency in North-western Ontario that services Ojibway, Cree, and Oji-Cree children, adolescents, and their families. The non-First Nations children and adolescents were obtained from the original authors of the Revised Ontario Child Health Study, and included children and adolescents from two mental health agencies and the general population from an urban area. Differential item and test functioning analyses were conducted on the following mental health scales: Regulation of Attention, Impulsivity, and Activity Level, Cooperativeness with Others, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood. The results indicated that differential item functioning was evident on most items of the mental health scales. However, the results from the differential test functioning analyses suggest that the bias evident at the item level was not sustained when the items were aggregated at the scale level. Clinically, these findings suggest that the mental health scale scores on the BCFPI can be interpreted the same ways across the cultural groups.

Differential Item and Test Functioning of the Brief Child and Family Phone Interview in First Nations and non-First Nations Children and Adolescents

It is becoming the standard for clinical practice to be based upon empirically supported research findings (Hunsley, 2003; Hunsley, Lee, & Wood, 2003; Lilienfeld, Lynn, & Lohr, 2003). It is important for mental health clinicians to base their assessment procedures on scientific evidence (Hunsley, 2003; Lilienfeld et al., 2003). Mental health service providers are often challenged with conducting clinical assessments with culturally diverse clients (Tseng, 2001). However, there is a lack of scientific evidence regarding clinical assessments among culturally diverse clients. The aim of the present study is to examine cultural differences between First Nations and non-First Nations children and adolescents on the Brief Child and Family Phone Interview (BCFPI). The BCFPI assesses emotional and behavioural symptoms and empirical investigation is required to examine whether these symptoms can be assessed in an accurate and unbiased manner in First Nations children and adolescents. An overview of Canada's First Nations population will be presented. The discussion will then be directed to the appropriateness of utilizing conventional clinical assessments with First Nations individuals. Finally, the modern psychometric techniques that will be applied to examine response differences will be illustrated.

The cultural differences between First Nations and non-First Nations may have a significant impact on the clinical assessment of psychopathology. Culture may influence the manifestations of symptoms and treatment of psychopathology (Cuellar, 2000). Marsella and Yamada (2000) define culture as:

Shared learned meanings and behaviours that are transmitted from within a social activity context for purposes of promoting individual/societal adjustment, growth and development. Culture has both external (i.e., artefacts, roles, activity contexts, institutions) and internal (i.e., values, beliefs, attitudes, activity contexts, patterns of consciousness, personality styles, epistemology) representations. The shared meanings and behaviours are subject to continuous change and modification in response to changing and external circumstances (p. 12).

In addition, Marsella and Yamada (2000) state that “culture is the lens or template we use in constructing, defining and interpreting reality ...even mental disorders must vary across cultures because they cannot be separated from cultural experience” (p. 12). Culture is guided by social and historical contexts (Jenkins & Karno, 1992). There is a need for mental health practitioners to be aware of culture and the influences on mental health (Tseng, 2003).

Canada's First Nations People

Demographics of First Nations People

Canada's First Nations people represent approximately 1 million people, as 976,305 individuals reported they were North American Indian, Métis, or Inuit in the 2001 Canadian Census (Statistics Canada, 2001). The provinces of Ontario, British Columbia, and Manitoba have the highest population of First Nations people, respectively. The highest concentration of First Nations people is in the north and on the prairies (Statistics Canada, 2001). Canada has approximately 2284 reservations (Frideres, 1993). However, as reported in the Census,

approximately 70% of the First Nations population do not reside on reservations and nearly one-half of First Nations people live in urban areas (Statistics Canada, 2001). There is a large cultural diversity among the Canadian First Nations population as there are 596 bands that represent 11 different languages with more than 58 dialects (Frideres, 1993). However, only one-quarter of First Nations people are capable of having a conversation in a First Nations language (Statistics Canada, 2001).

The First Nations population is increasing. The average age of the First Nations population is younger than the average age of the general population. Although the total First Nations population represents approximately 4% of the Canadian population, First Nations children under age 15 represent 33% of the First Nations population, compared to 19% of the general population. In addition, First Nations children are more likely to live in a lone-parent family rather than living with both parents, this is twice the rate within the general population (Statistics Canada, 2001).

Historical and Contemporary Issues of First Nations People

As mentioned previously, culture is grounded in social and historical contexts (Jenkins & Karno, 1992). Barwick, Schmidt, and Hodges (2004) note that for clinicians to effectively serve First Nations clients, clinicians must understand the historical and contemporary issues that are relevant to First Nations people in Canada. A number of First Nations people are likely to be impacted by social factors such as limited educational opportunities, poverty, access to health care, substandard housing, and have poor sanitation and water quality (Royal Commission on First Nations People, 1995). In addition, First

Nations people are also likely to face significant economic issues such as unemployment, access to employment, and racism in the work place (Royal Commissions on First Nations Peoples, 1993). Historically, First Nations people have had many political hardships, as the policies developed by the Canadian government were attempts to suppress and eradicate culture among First Nations people and were intended to serve the political and economic interests of the dominant society (Titley, 1986).

During the confederation in 1867, the federal government under the British North American Act, assumed responsibility over First Nations people. In 1876, the Indian Act was established and aimed to suppress culture and traditional ceremonies. First Nations people were prohibited from performing religious and cultural ceremonies until 1951 when an amendment was made to the Indian Act. In addition, provisions under the Indian Act also lead to over 100,000 First Nations people losing their identification of Status Indians or Registered Indians in Canada (Status Indians have advantages over Non-Status Indians or Métis such as tax exemption and on reserve housing). First Nations people lost their status by giving up their status in exchange for the right to vote in the federal election otherwise First Nations people were not permitted to vote. First Nations women lost their status when they married non-First Nations men, and any offspring were not recognized as having status. It was not until 1985, with the introduction of Bill C-31, that people could apply to have their status reinstated (Barwick et al., 2004).

The Indian Act gave the federal government responsibility for education of First Nations children. Residential schools were government-mandated from 1879

to as late as 1973. By the 1940s approximately half of the First Nations student population was enrolled in residential schools across Canada (Assembly of First Nations, 1994; Miller, 1996). The Christian church and Canadian government believed that First Nations people were uncivilized and primitive and needed to be “civilized” into mainstream society. The rationale for residential schools was to assimilate First Nations children to dominant customs, to segregate First Nations children from their families and heritage, and integrate First Nations children into mainstream society (Assembly of First Nations, 1994). “Schooling meant more than just teaching skills. It was to be the major instrument for the destruction of First Nations culture and for grooming the First Nations population for a place on the lower rung of the social ladder” (Titley, 1986, p. 1814). First Nations children were called savages, taught their traditional customs and ceremonies were evil and were forbidden from speaking their First Nations language (Assembly of First Nations, 1994). This attempt at cultural genocide left First Nations students with emotional, mental, physical, spiritual, and sexual abuse (Assembly of First Nations, 1994; Haig-Brown, 1988). Consequently, the effects of residential schools have been long lasting and multi-generational. The aftermath of residential schools continues today and remains a significant issue faced by First Nations people (Barwick et al., 2004). These short-sighted government policies have caused widespread mental health problems among First Nations people (Kirmayer, Brass, & Tait, 2000).

Mental Health and First Nations People

There is a high prevalence rate of psychopathology among the First Nations population (Roy, Choudhuri, & Irvine, 1970). Researchers (Dalrymple,

O'Doherty, & Nietschei, 1995) found that First Nations individuals are 33% more likely than non-First Nations individuals to be admitted to acute psychiatric care and are likely to remain in psychiatric care twice as long as non-First Nations individuals. Epidemiological studies provide important information about psychopathology. However, few epidemiological studies have examined the prevalence of mental disorders among First Nations people (Dillard & Manson, 2000). Kirmayer et al. (2000) have indicated that four epidemiological studies have been conducted and only two of these studies have been conducted in Canada. A lot of research on First Nations mental health originates from the United States. Gotowiec and Beiser (1993) indicate that generalizations made to Canada's First Nations people must be done cautiously as there are different histories between the countries. It is also important to note, it is not the intention of this report to generalize or label First Nations adults or children but to present the common mental health issues that First Nations people are challenged with.

Roy and colleagues (Roy et al., 1970) examined the prevalence rate of psychiatric disorders from hospital admissions among 10 First Nations reservations (1 Salteaux and 9 Cree) in Northwest Saskatchewan. Non-First Nations individuals from 18 surrounding rural communities were used as comparisons for the study. The clinician based the diagnoses upon the ICD criteria. The First Nations communities revealed a significantly higher rate of mental disorders than the surrounding non-First Nations communities. The mental disorders included: neurosis (3.2%), mental deficiency (2.5%), schizophrenia (2.2%), functional psychoses (1.3%), and alcoholism (0.5%). The First Nations communities showed higher rates of schizophrenia and mental

deficiency than their non-First Nations neighbours. The researchers highlight that their methodology could not reveal an accurate prevalence rate of mental disorders. Moreover, alcohol use was prevalent, however, the results did not reveal the typical alcohol abuse patterns.

Sampath (1974) interviewed 214 Inuit adults from a Southern Baffin Island community with a population of 550. Through the assistance of a translator, the author found that 37% of the adults interviewed met the DSM-II criteria for a mental disorder. The prevalence rate of the mental disorders included: neuroses (11.6%), affective disorders (4.6%), psychoses (3.8%), personality disorders (2.8%), and schizophrenia (0.5%). Sampath (1974) found that on average, women reported more psychiatric symptoms than men.

The prevalence of psychiatric disorders has also been studied from researchers across the border. Shore and colleagues (Shore, Kinzie, Hampson, & Pattison, 1973) assessed psychiatric disorders among adults from a Pacific Northwest First Nations village. The 100 First Nations individuals were from an isolated fishing and lumbering community that had a total population of 500 people. The psychiatrists conducted clinical interviews with the DSM-I and found mental disorders of: alcoholism (27%), psychoneurosis (18%), psychophysiologic reactions (9%), and personality disorders (5%). Alcoholism was seen as the primary psychiatric disorder and the main health issue in the community. Sex was an important indicator of type of mental illness, as men were more likely to report alcoholic tendencies and women were more likely to report psychoneurotic symptoms.

Most recently, Kinzie and colleagues (Kinzie, Leung, Boehnlein, Matsunaga, et al., 1992) examined the prevalence of psychiatric disorders among First Nations clients. These researchers conducted a follow up study to the First Nations from the American North West Coast community originally interviewed by Shore and colleagues (Shore et al., 1973). This study utilized the Schedule for Affective Disorders and Schizophrenia Lifetime version interviews. The participants were matched with the original sample for sex and age. In total, 31.4% of the First Nations met the DSM-III-R criteria for a mental disorder. Most of the individuals who were employed did not meet the criteria for the diagnosis of a mental disorder. Moreover, the variables age, marital status, and educational level were not associated with the diagnosis of a psychiatric disorder. The point prevalence and lifetime prevalence, respectively, included the following mental disorders: alcohol abuse and dependence (20.9%, 56.9%), substance abuse (1.0%, 3.5%), affective disorders (7.0%, 28.0%), dementia (1.1%, 1.1%), organic anxiety (0.4%, 0.4%), organic personality (2.0%, 2.0%), schizophrenia (2.1%, 2.1%), simple phobias (0.6%, 0.6%), panic without agoraphobia (0.6%, 0.6%), post-traumatic stress disorder (2.1%, 5.0%), and personality disorders (1.4%, 1.4%). In comparison to the original study, Shore and colleagues (Shore et al., 1973) found that 59% of the participants met the criteria for a mental disorder while Kinzie and colleagues (Kinzie et al., 1992) found that 31.4% met the criteria for a mental disorder. In addition, in both studies, alcohol-related disorders are prominent. These service utilization and epidemiological studies only provide a crude indication of the prevalence and type of mental disorders facing First Nations clients. As mentioned previously, these prevalence rates are likely to not provide

an accurate estimate of the incidence of psychiatric disorders among First Nations. The actual prevalence rate of psychiatric disorders may be higher, as many First Nations people do not obtain treatment (Kirmayer et al., 2000).

Research regarding the mental health of First Nations children is scarce (Kirmayer et al., 2000). Beals et al. (1997) examined the prevalence of psychiatric disorders among Northern Plain First Nations adolescents who had participated in an earlier research study. The Diagnostic Interview Schedule for Children Version 2.1C was used to generate Diagnostic Statistical Manual-III-Revised (DSM-III-R) diagnoses. Twenty-nine percent of the 109 Northern Plain First Nations adolescents met criteria for at least one DSM-III-R diagnosis. The results for the diagnostic categories included substance use disorders (18.3%), disruptive behaviour disorders (13.8%), anxiety disorders (5.5%), affective disorders (4.6%), and eating disorders (1%). The most common diagnoses were alcohol dependence/abuse (11.0%), attention deficit hyperactivity disorder, marijuana dependence/abuse (8.6%), major depressive disorder (4.7%), and other substance/abuse (3.9%). There was a high rate of comorbidity among this sample, as 53.3% of those diagnosed with a disruptive disorder also met diagnostic criteria for a substance use disorder. Sixty percent of the adolescents diagnosed with a depressive disorder were also diagnosed with a substance use disorder.

It is recognized that First Nations children are more at risk to develop mental health problems than the general population (Gotowiec & Beiser, 1993). American First Nations children are 20% to 25% likely to experience an emotional disorder, compared to 5% to 15% of non-First Nations children (Yates,

1987). In 1997, the First Nations and Inuit Regional Health Survey was conducted across Canada inquiring about the health of First Nations and Inuit children and adolescents. The survey results indicate that overall 17% of the children have an emotional or behavioural problem. The prevalence of emotional and behavioural problems increased with age, with 9% in the newborn to 5-year age group, 18% in the 6 to 11 year age group, and 23% in the 12-17 age group (MacMillan, Walsh, Jamieson, Crawford, & Boyle, 2000).

Despite the acknowledgement that many First Nations children and adolescents exhibit mental health issues, many people are not aware that First Nations children have different mental health concerns than their dominant culture counterparts. Suicide and substance abuse are significant mental health concerns among First Nations children and adolescents (Gotowiec & Beiser, 1993). Suicide is a critical problem facing First Nations people (Gotowiec & Beiser, 1993; Royal Commission on First Nations People, 1995). The Royal Commission on First Nations Peoples estimated that the suicide rate for First Nations individuals was approximately three times higher than the general population (Royal Commission on First Nations People, 1995). Adolescents and young adults are most at risk, as the suicide rate for First Nations individuals between the ages of 10 and 19 years are five to six times higher than the general population (Royal Commission on First Nations People, 1995). Suicide among First Nations individuals is less likely to be related to mental illness such as depression, anxiety disorders, or schizophrenia and is more likely to be related to alcohol use, brain damage, paranoid psychosis from solvent use, socio-economic

factors, or cultural stress (e.g., racial discrimination, suppression of cultural beliefs and spirituality) (Royal Commission on First Nations People, 1995).

Substance abuse is a prominent mental health risk among First Nations children and youth (Gotowiec & Beiser, 1993). Lalinec-Michaud, Subak, Ghadirian, and Kovess (1991) conducted a survey in 1985 in rural Quebec to examine the use of alcohol, tobacco, cannabis, and other illicit substances among First Nations and Francophone high school students. The First Nations adolescent group reported higher use of amphetamine, hallucinogen, and inhalant use than the Francophone group. In terms of alcohol and tobacco use, there were no significant differences between the First Nations and Francophone groups, except that the Francophone students reported more regular consumption of alcohol. Similarly, another Canadian survey found that First Nations adolescents reported more use of substances such as cannabis, LSD, solvents, and other hallucinogens than their non-First Nations dominant culture counterparts (Gfellner & Hundleby, 1995). In addition, adolescent First Nations females in Ontario have also reported utilizing non-prescriptions drugs more than any other age group (Myers, Bullock, Calzavara, Cockerill, & Marshall, 1997).

Culture and Clinical Assessment

It is not only important for clinicians to be clinically competent, it is also important for clinicians to be culturally competent (Tseng, 2003). Cultural competency has been suggested as a requirement when providing mental health services to clients of diverse cultural backgrounds (Tseng, 2001). There is a consensus that it is essential to acknowledge cultural factors during clinical

assessments (Butcher, Nezami, & Exner, 1998; Dillard & Manson, 2000; Lopez, 2000; Lu, Lim, & Mezzich, 1995; Tseng, 2001, 2003; Van de Vijver, 2000). This is essential as cultural factors are likely to heavily influence the clinical assessment (Rogler, 1993). The clinician should acknowledge cultural factors in the assessment process and the influences these culture factors have on the clinical assessment (Tseng, 2001). The clinical assessment evaluates the client's behaviour, problems and aids in treatment formulation. Hence, it is imperative that the clinical assessment be an accurate representation of the client's psychopathology. The clinician needs to recognize how these cultural factors influence psychopathology. The client's perception, experience, expression, and presentation of their psychopathology or symptom manifestations are impacted by cultural factors and this would affect the clinical assessment (Tseng, 2001, 2003).

First Nations mental health resources are typically confined to urban mental health services (Royal Commission on Aboriginal People, 1992). Culturally diverse clients are typically assessed with traditional assessment measures (Pollack & Shore, 1980). Assessments are often plagued with complexity and ambiguity and involve special considerations to ensure an accurate assessment of culturally diverse clients (Butcher et al., 1998). Most assessment instruments of psychopathology have been constructed by Caucasian researchers and clinicians and have been developed for the general population. This elicits concern about the value and appropriateness of applying conventional assessment instruments to individuals with different cultural backgrounds, as the assessment instruments' reliability, validity, and applicability are called into

question (Tseng, 2001, 2003). Despite the fact that culturally diverse clients have different cultural backgrounds, extremely critical decisions and treatment formulations are based upon clinical assessment instruments that have been devised for the general population (Butcher et al., 1998).

Psychological assessments of culturally diverse groups have been questioned due to the considerable cultural differences in test responses. This assessment procedure typically encompasses conventional psychological assessments that have been standardized among the general population (Pollack & Shore, 1980). Shore and Manson (1981) note that First Nations test responses may differ from the typical non-First Nations norms. The standardized norms are likely to result in invalid comparisons (Bonder, Martin, & Miracle, 2002), as First Nations individuals are not likely to be included in the standardization of norms (Dillard & Manson, 2000). Researchers have found that adult First Nations participants reported substantially elevated scores on depression, hypomania, hypochondriasis, hysteria, schizophrenia, paranoia, psychopathic deviate, psychasthenia, and social introversion when compared to their dominant culture counterpart on instruments of psychopathology. Researchers have contested that these psychological profiles are substantially influenced by cultural variations in test response (Borzecki, Wormith, & Black, 1984; Butcher, Braswell, & Raney 1983; Graham, 1993; Herreid & Herreid, 1966; Hoffman, Dana, Bolton, 1985; Kline, Rozyrko, Flint, & Roberts, 1973; Mandelzys & Lane; 1980; Page & Bozlee, 1982; Panton, 1980; Pollack & Shore, 1980; Uecker, Boutilier, & Richardson, 1980; Venn, 1988). Consequently, it is difficult to conclude if elevated symptomatology is an indication of psychopathology or cultural differences (Dillard & Manson, 2000). The available

research literature poses important questions regarding the utility of these conventional assessment instruments with culturally diverse clients (Allen, 1998). Subsequently, there is a proliferation of dissatisfaction with conventional assessment instruments of psychopathology being utilized for culturally diverse clients (Hoffman, Dana, & Bolton, 1985).

Brief Child and Family Phone Interview

In 2001, the Ministry of Community and Social Services mandated that all Ontario funded providers of Children Mental Health Services utilize the Brief Child and Family Phone Interview (BCFPI; Cunningham, Pettingill, & Boyle, 2001). The BCFPI is a structured interview that is administered to parents, teachers or adolescents. The BCFPI can be administered in 30 to 45 minutes by either phone or in person. The BCFPI is administered at intake and is completed before clinical assessment and treatment is conducted. The BCFPI assesses emotional and behavioural disorders in children and adolescents. The structured interview evaluates Regulation of Attention, Impulsivity, and Activity Level, Cooperation with Others, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood. In addition, the BCFPI also explores how these problems influence the child's academic, social, and family's functioning. Moreover, the BCFPI also examines whether the interviewee would be interested in attaining additional resources (e.g., literature, workshops, courses, parent support groups) and whether there would be any barriers (e.g., transportation, child care, work schedules) that may interfere with their ability to access services (Cunningham et al., 2001).

The BCFPI Regulation of Attention, Impulsivity and Activity Level scale contains six-items that assess the child's ability to concentrate, finish tasks, avoid distractions, and control activity and impulsive responding. For instance, an item from the Regulation of Attention, Impulsivity, and Activity Level scale includes "Has difficulty following directions or instructions." High scale scores indicate the child has difficulty regulating activity level and impulsivity. These traits are consistent with Attention Deficit Hyperactivity Disorder symptoms as outlined in the DSM-IV (Cunningham et al., 2001).

The BCFPI Cooperativeness with Others scale contains six-items that examine how well the child cooperates with other people. For example, an item from the Cooperativeness with Others scale includes "Defiant, talks back to adults." High scale scores indicate defiant and noncompliant behaviours with peers and adults that are consistent with symptoms of Oppositional Defiant Disorder in the DSM-IV (Cunningham et al., 2001).

The BCFPI Conduct scale contains six-items that assess for antisocial behaviour and the extent to which the child obeys rules and regulations. For instance, an item from the Conduct scale includes "Has broken into a house, building or car." The endorsement of these items seldom occurs among the general population; consequently, high scale scores will occur when a few items on the scale are highly endorsed or when several items are low endorsed. These items are consistent with the symptoms of Conduct Disorder as described in the DSM-IV (Cunningham et al., 2001).

The BCFPI Separation from Parents scale contains six-items that examine the child's ability to separate from their parents. The degree of separation

difficulty is assessed with how much the child worries and how upset the child becomes. For example, an item from the Separation from Parents scale includes “Worries about being separated from loved ones.” High scale scores indicate that the child experiences difficulty when being separated from his/her parents. This scale corresponds to the DSM-IV’s Separation Anxiety Disorder (Cunningham et al., 2001).

The BCFPI Managing Anxiety scale contains six-items that assess the child’s anxiety or worries about performance, past and present behaviour, and future events. For instance, an item from the Managing Anxiety scale includes “Worries about doing better at things.” High scale scores are consistent with Anxiety Disorders in the DSM-IV (Cunningham et al., 2001).

The BCFPI Managing Mood scale contains nine-items that examine the child’s mood, anhedonia, and suicidal ideation. For example, an item from the Managing Mood scale includes “Gets no pleasure from usual activities” and “Deliberately harms or attempts suicide.” High scale scores correspond to symptoms of Major Depression as outlined in the DSM-IV (Cunningham et al., 2001).

The BCFPI Child and Family Functional Impact scales are indicative of how the child’s emotional and behaviour symptoms are impacting or influencing the child’s social, academic, and family functioning. The Child’s Social Participation scale contains three-items that assess the child’s participation in social relations. For instance, an item from the Child’s Social Participation scale includes “Doing things less with other kids.” High scale scores suggest that the

child may be isolating or withdrawing from spending time with peers

(Cunningham et al., 2001).

The BCFPI Quality of the Child's Social Relationships scale contains three-items that examine the child's relationships with parents, peers, and teachers.

For example, an item on this scale includes "Getting along with you and your partner." High scale scores indicate that the child maybe experiencing poor relationships (Cunningham et al., 2001).

The BCFPI School Participation and Achievement scale contains three-items that assess the child's academic functioning, attendance, and relationship with the teacher. For instance, an item on this scale includes "Getting along with his/her teachers." High scale scores are indicative of poor academic functioning, poor attendance, and or a poor relationship with the teacher (Cunningham et al., 2001).

The BCFPI Family Activities scale contains four-items that examine the degree to which the child's emotional and behaviour difficulties have impacted the family activities and functioning. For example, an item on the Family Activities scale includes "How frequently has XX's behaviour's prevented his/her brothers or sisters from having friends, relatives or neighbours to your home?" High scores indicate that the child's behaviour has limited the family's social networks and functioning (Cunningham et al., 2001).

The BCFPI Family Comfort scales contains three-items that assess the degree to which the child's emotional and behavioural difficulties have impacted the internal family functioning. For instance, an item from this scale includes "How frequently has (child's name) behaviour caused you to be anxious or

worried about his/her chances for doing well in the future?" High scale scores suggest the child behaviour is influencing the anxiety level or comfort within the family network (Cunningham et al., 2001).

The BCFPI generate composite scales for externalizing behaviour based on the Regulation of Attention, Impulsivity, and Activity Level, Cooperativeness with Others, and Conduct scales. The BCFPI also generates composite scales for internalizing behaviour based on Separation from Parents, Managing Anxiety, and Managing Mood scales. A Total Problem scale is also generated by the BCFPI and is based on the Externalizing Problem scale, Internalizing Problem scale, Child Functional Impact Scale, and Family Functional Impact Scale.

The items for the BCFPI were selected from the Revised Ontario Child Health Study (Boyle et al., 1993a; Boyle et al., 1993b; Cunningham et al., 2001). The original Ontario Child Health Study was conducted in 1983, in which the authors developed a problem checklist of emotional and behaviour symptoms outlined in the DSM-III (Boyle et al., 1987). The Ontario Child Health Study scales were revised to correspond to the DSM-III-R for childhood disorders to assess the prevalence of emotional and behavioural disorders in children aged 6 to 16 years. The authors developed a problem checklist to evaluate attention-deficit hyperactivity disorder, oppositional disorder, conduct disorder, separation anxiety, overanxious disorder, and depression. The problem checklist was administered to parents, teachers, and youth within the general population and from two child mental health agencies (Boyle et al., 1993).

The BCFPI is an abbreviated version of the Revised Ontario Child Health Study. The Revised Ontario Child Health Study contained 14-items to assess

attention-deficit hyperactivity disorder and the BCFPI contained 6-items to evaluate regulation of attention, impulsivity, and activity level; the Revised Ontario Child Health Study contained 9-items to assess oppositional defiant symptoms and the cooperativeness with others of the BCFPI contained 6-items; the Revised Ontario Child Health Study contained 13-items to assess conduct disorder symptoms and the conduct scale of the BCFPI contained 6-items; the Revised Ontario Child Health Study contained 9-items to assess separation anxiety disorder symptoms and the separation from parents scale of the BCFPI contained 6-items; the Revised Ontario Child Health Study contained 7-items to assess overanxious symptoms and the managing anxiety scale of the BCFPI contained 6-items; the Revised Ontario Child Health Study contained 16-items to assess depression and the managing mood scale of the BCFPI contained 9-items (Boyle et al., 1993; Cunningham et al., 2001). Cunningham and colleagues (2001) selected the items from the Revised Ontario Child Health Study that provided the most accurate assessment of each subscale for the BCFPI. Furthermore the BCFPI corresponds to emotional and behavioural symptoms described in the DSM-IV criteria. Although the questions from BCFPI are based upon diagnostic criteria, the authors indicate that the BCFPI is not a diagnostic tool (Cunningham et al., 2001).

The BCFPI has clinical norms and community sample norms that are based upon the responses from the Revised Ontario Child Health Study (Cunningham et al., 2001). Boyle et al. (1993a) indicates the clinical norms were selected from two mental health agencies and the community sample was selected among the general population. The Revised Ontario Child Health Study

does not indicate whether First Nations children and adolescents were included in the study. However, the original Ontario Child Health Study explicitly indicates that First Nations children were excluded from the study (Boyle et al., 1987). The BCFPI has been exposed to high scientific scrutiny among the general population, in which the BCFPI has proved to be a psychometrically sound and clinically useful measurement of emotional and behavioural problems in children and adolescents. However, these high standards have not been established for First Nations clients.

As mentioned previously, First Nations test responses may differ from the typical non-First Nations norms. Comparing First Nations test responses to standardized norms based upon the general population may result in invalid comparisons as First Nations individuals are not likely to be included in the standardization of norms (Bonder et al., 2002; Dillard & Manson, 2000; Shore & Manson, 1981). Hence, the purpose of the present study was to examine whether the BCFPI is an appropriate assessment of psychopathology among First Nations children and adolescents and to understand the impact that culture has on this clinical assessment instrument. To the authors' knowledge, no previous research has examined the cultural impact of First Nations children and adolescents on the BCFPI. Further research is warranted as the BCFPI is mandated across Ontario. As mental health care providers, Tseng (2001) indicates that it is important to learn how culture influences the assessment of psychopathology.

Item Response Theory

Fundamentals of Item Response Theory

Item response theory is considered one of the most promising psychometric techniques (Bejar, 1983; Embretson & Reise, 2000; Santor & Ramsay, 1998). Item response theory was developed to explain and analyze the relationship between characteristics of the individual and responses to individual test items (Lord, 1980). Item response theory proposes that performance on a test can be explained or predicted by the individual's abilities, traits, or attributes. This model utilizes estimate scores of these attributes to explain or predict performance. These characteristics are referred to as latent traits or abilities as the attributes cannot be measured directly (Lord & Novick, 1968). Lord and Novick (1968) mathematically represent latent traits as

$$\theta = (\theta_1, \theta_2, \dots, \theta_k).$$

Item response theory postulates that the unobservable traits or abilities are represented by the observable item and test performance. The item response model theorizes that the individual's unobservable traits are accountable for observable item performance (Hambleton & Swaminathan, 1985). Psychological constructs are conceptualized as latent traits that account for human behaviour, and these psychological constructs are measured indirectly, as the latent attributes are measured by responses to test items (Embretson & Reise, 2000). There needs to be a clear relationship between the attributes the test is designed to measure and the individual's responses to the items on the test. For instance, in a well-designed measure of depression, individuals who exhibit depressive symptomatology would be more inclined to endorse depressive symptoms

(Murphy & Davidshofer, 1998). Santor and Ramsay (1998) indicate that endorsement of an item may be attributed to three factors (1) the individual's trait or ability, (2) the potential of the item to represent the trait or ability, and (3) the population in which the individual is selected.

Item response theory begins with a specific set of assumptions based upon the test data regarding the mathematical relationship between an individual's trait or ability and the probability of endorsing a test item (Hambleton & Swaminathan, 1985; Murphy & Davidshofer, 1998). The assumptions impose specific statistical constraints that declare that the analyses are valid (Bejar, 1983). These key assumptions of item response theory are dimensionality and local independence.

Dimensionality of latent traits. A fundamental assumption to understanding item response theory is the dimensionality of latent traits. A unidimensional model is considered when only a single latent attribute is presumed to result in the item endorsements or to explain test performance (Hambleton & Swaminathan, 1985; Reise, 1999). However, Hambleton and Swaminathan (1985) note that this assumption cannot be strictly adhered to, as there are always other personality, cognitive, and test-taking factors that influence test performance. To adequately satisfy this assumption, the test data set should consist of a dominant factor that represents the trait measured by the test. Item response models that consist of more than one single trait are considered to be multidimensional (Hambleton & Swaminathan, 1985).

Hambleton and Swaminathan (1985) indicate that the unidimensionality of a test may vary within populations, as a test may be unidimensional in one

population and have more than one dimension in another population. For instance, a test may have a high cultural component and may be unidimensional for those individuals with a similar cultural background and may not be unidimensional to individuals with different cultural backgrounds (Hambleton & Swaminathan, 1985).

Local independence. Another fundamental assumption of item response theory is local independence. When numerous items are examined it is important to consider the assumption of local independence to ensure the data is suitable for the theoretical item response model (Embretson & Reise, 2000; Lord & Novick, 1968). “*Local independence means that within any group of examinees all characterized by the same values $\theta_1, \theta_2, \dots, \theta_k$, the (conditional) distributions of the item scores are all independent of each other*” (Lord & Norvick, 1968, p. 361). This implies that the item scores are associated by the latent variables values $\theta_1, \theta_2, \dots, \theta_k$ (Lord & Norvick, 1968). The assumption of local independence diverges from classical theory of internal consistency, as the traditional theory favours high correlations among items. In contrast, this assumption contests that an individual’s test item responses to the various test items are not correlated and are statistically independent of each other. Local independence is attained when the item response model differentiates between items (Embretson & Reise, 2000).

Local independence is associated with the number of different attributes that influence item responses (Embretson & Reise, 2000). There is a relationship between local independence and unidimensionality. Local independence provides verification that the assumption of unidimensionality has been

established if the item response model has one dimension (Embretson & Reise, 2000; Bejar, 1983). As Lord (1980, p. 19) explains “local independence follows automatically from unidimensionality.” Hence, Bejar (1983) contests that in illustrating that item responses are unidimensional, it can be assumed that the assumption of local independence has been met. Moreover, it is also possible for local independence to be attained for multidimensional information, providing the item response model has data for each dimension (Embretson & Reise, 2000).

Item Response Theory and Item Characteristic Curves

Item response models are non-linear mathematical functions that were constructed specifically for the purpose of comprehending how individual differences in attributes influence behaviour when responding to a test item. Item response theory proposes that item characteristic curves are a class of logistic functions that represent a simple relationship between the individual's attribute and the probability of the individual endorsing test items (Hambleton & Swaminathan, 1985; Murphy & Davidshofer, 1998; Reise, 1999). The item characteristic curves are based upon these assumptions and represent an amalgamation of assumptions regarding fundamental relationships and the empirical results. Providing these assumptions are accurate, item response models allow for accurate conclusions to be made regarding attributes from responses to test items. This psychometric model allows researchers to determine why people endorse items (Murphy & Davidshofer, 1998). Hence, individuals with elevated traits are likely to endorse items that reflect those traits. The observable and unobservable relationship is accounted for by a mathematical

function (Embretson & Reise, 2000; Hambleton & Swaminathan, 1985; Meier, 1994).

An item characteristic curve is a non-linear regression function of item responses on the measured attribute. When using item response models it is necessary to select one of many mathematical functions to work as the form of the item characteristic curves. This provides the opportunity for verification of how well the selected model accounts for test results. In situations where only one attribute is measured, the regression is referred to as an item characteristic curve. The term item characteristic function is used when more than one attribute is measured. The curve attaching the means of the distributions consists of the regression of item score on the measured attribute (Hambleton & Swaminathan, 1985).

Item characteristic curves are important in differentiating between non-discriminate and discriminate test items (Meier, 1994). For instance, on a measure that assesses depression, an adequate discriminating item would differentiate between the depressed clients and the non-depressed clients. Meier (1994) indicates that is important to ensure precise measurement of an attribute by identifying inadequate discriminating items and eliminating these items from the measurement.

The item characteristic curves depict most of the information from the item analyses and shows how this information may be utilized to comprehend the relationship between item responses and the measured attribute (Lord, 1980). The item characteristic curve provides an illustrative summary of the essential features of a test item (Murphy & Davidshofer, 1999). The item characteristic

curves are defined when the general form is specified and the curve parameters of an item are identified and these are graphically depicted (Hambleton & Swaminathan, 1985). Figure 1 illustrates an example of item characteristic curves. Note that the x-axis represents the trait levels and that the y-axis represents the item response probability. Embretson and Reise (2000) note that item characteristic curves can vary and the item responses may differ on slope. The slopes of the item characteristic curve will vary to depict the differences in test items (Hambleton & Swaminathan, 1985). The slope represents the item discriminations and depicts the probability of change at the attribute level (Embretson & Reise, 2000). The steepness of the slope is important as it discriminates those individuals who possess the attribute from those individual who do not. The steeper the upward slope indicates the individuals who possess higher levels of the attribute are more likely to endorse that item (see Figure 1). A flatter slope or straighter line discriminates evenly at all levels and the individuals did not possess a significant amount of the attribute that was measured (see Figure 1). The downward slope indicates that individuals who possess higher levels of the attribute are not likely to endorse that item (see Figure 1) (Murphy & Davidshofer, 1998). In addition, the form of an item characteristic curve is an indication of the changes in the attribute and conveys the changes in the probability of endorsing a certain response at various points along the latent trait continuum. For dichotomous items (i.e., true/false questions), when a certain response is judged to be in “agreement” with an item, the item characteristic curve regresses the probability of item endorsement at the attribute level. For polytomous items (i.e., Likert rating scales), the item characteristic

curve regresses the probability of item endorsement in each category at the attribute level (Embretson & Reise, 2000). The endorsement of trait levels is often represented as z-scores that have a mean of 0.00 and a standard deviation of 1.00. However, other metric representations are possible and can be applied to IRT analyses (Waller, Thompson, & Wenk, 2000).

Item response theories are acknowledged as a “strong” modeling method as the fulfillment of certain assumptions is required (Embretson & Reise, 2000). Bejar (1983) states that item response models can account for performance at the item level. However, to utilize the power of item response theory and to obtain an accurate estimate of the attributes being measured a large sample size is essential (Bejar, 1983). Ironson (1983) recommends that large sample sizes are required to establish precise estimations for the different parameters.

Differential Functioning

Item response models are designed to examine whether extraneous characteristics such as sex, age, or ethnicity may influence the assessment of psychological or cognitive functioning (Camilli & Shepard, 1994). Different groups of individuals (e.g., First Nations vs. non-First Nations) may respond differently to the same test question (Holland & Thayer, 1987). It is important to understand how an individual’s background characteristics influence the assessment process, as this is pivotal for diagnosis and treatment planning (Camilli & Shepard, 1994). Two different groups of individuals can be compared at the latent level and assessed to distinguish whether the groups differ on the response to the test items or the total scale score (Raju et al., 1995). Item

response models refer to this as differential item functioning and differential test functioning.

Differential item functioning. Waller et al. (2000) indicate that differential item functioning is a psychometric term, however, item bias is the judgment placed on the item regarding the ethical, personal, or social implications of differential item functioning. In the present study, cultural bias is the social implications of differential item functioning. It is important to note that the terms differential item functioning and item bias are used interchangeably throughout the discussion. Santor and Ramsay (1998) note that the detection of differential item functioning is a fundamental psychometric issue to ascertain that attributes are assessed impartially and fairly by a measure. Waller et al. (2000) indicate that differences on items or test scale scores may be attributed to (1) test bias, (2) actual differences between the groups, and (3) a combination of both. Ironson (1983) notes that item response theory presents advancement in establishing differential item functioning, as these techniques are less likely to falsely identify an item as biased when compared to the traditional or classical test theory procedures. Due to the strong theoretical foundation of item response theory, it is considered one of the most promising psychometric techniques at identifying differential item functioning. Differential item functioning can be defined as “individuals from different populations, but who have the same total score on the test, have different probabilities of responding correctly to that item” (Hambleton & Swaminathan, 1985, p. 284). Differential item functioning is detected when individuals who report equal symptomatology, respond to items differently. Differential item functioning is

identified when the likelihood of responding to an item for one group, differs from the likelihood of responding to that item from another group (Ironson, 1983; Lord, 1980; Millsap & Everson, 1993; Raju, van der Linden, & Fler, 1995; Reise, 1999; Santor & Coyne, 2001; Santor & Ramsay, 1998, Santor et al., 1994). Differential item functioning may be attributed to the test item measuring the attribute differently for different groups of individuals (Ironson, 1983).

Measurement between groups is unbiased when the item characteristic curves for two groups do not differ. When the item characteristic curves are similar, this suggests that two groups who are matched on the latent trait have the same probability of endorsing the item similarly. This suggests that the test measured the trait equally among the two groups. For illustration purposes, consider the two-item characteristic curves illustrated in Figure 2. The item characteristic curves represent the responses for First Nations and non-First Nations responses to the item “Has difficulty following directions or instructions” from the Regulation Attention, Impulsivity, and Activity Level scale from the BCFPI. The solid line represents the item characteristic curve for the First Nations individuals and the dashed line represents the item characteristic curve for the non-First Nations individuals. Note that the item characteristic curve is similar and overlaps at most levels of the latent trait (Ironson, 1983, Reise, 1999; Waller et al., 2000). Again for illustration purposes, consider the item characteristic curves illustrated in Figure 3 for the item “Has difficulty following directions or instructions.” The item characteristic curve for the First Nations individuals is higher than the item characteristic curve for the non-First Nations individuals at all levels of the latent trait. This represents uniform differential

item functioning, as the item characteristic curves differ between the two ethnic groups (Hambleton, Swaminathan, & Rogers, 1991). It is important to note that the variation or space between the item characteristic curves varies across the different levels of the latent trait values. Lastly, consider the item characteristic curves in which the curves between the two groups are not uniform among the trait levels (Hambleton et al., 1991). Note Figure 4, which illustrates a hypothetical example of non-uniform differential item functioning for the “Has difficulty following directions or instructions” item. The item characteristic curves vary between the groups at the low trait level values and the high trait level values. At the low trait level values the non-First Nations individuals scored higher than the First Nations individuals. However, at the high trait level values the First Nations individuals scored higher than the non-First Nations individuals. When non-uniform differential item functioning is evident, conclusions should be drawn cautiously (Waller et al., 2000). To summarize, when there is a difference between item characteristic curves, the item functions differently between the groups and additional investigation would be warranted (Ironson, 1983; Reise, 1999).

It is also important to consider the dimensionality of the traits when examining differential item functioning or item bias. For differential item functioning analyses to be accurate the observed scores should be constrained to assessing the unobservable traits the scale is intended to measure. Extraneous latent variables may influence the observed scores in an unexpected manner. Consequently, the extraneous latent variables may contribute to differential item functioning (Millsap & Everson, 1993).

Differential test functioning. Differential test functioning is similar to differential item functioning, except that differential test functioning examines bias at the test or total scale level whereas differential item functioning examines bias at the item level. Differential test functioning utilizes the test characteristic curve in analyses. The test characteristic curve represents the non-linear regression similar to the item characteristic curve. The test characteristic curve is calculated by aggregating the item characteristic curves on all items of the scale. The test characteristic curve can be mathematically defined as the following:

$$T_i = \sum_{j=1}^J P_j(\theta_i)$$

T_i represents the predicted true score for subject i , J represents the number of items contained on the scale, P represents the probability the latent trait will be endorsed by an individual. Hence, the test characteristic curve characterizes the logistic function of the predicted true scores on the estimated latent trait values for the individuals (Hambleton et al., 1991; Waller et al., 2000). For illustrative purposes consider Figure 5, which depicts the test characteristic curves for the six-item scale Regulation of Attention, Impulsivity, and Activity Level from the BCFPI between First Nations and non-First Nations individuals. As the test characteristic curves are similar, estimated trait values between the two ethnic groups are similar at all levels of the latent trait, indicating no differences between the two groups. However, now consider Figure 6, which illustrates test characteristic curves from the same BCFPI scale. The test characteristic curves vary on levels of the latent trait or attention, impulsivity, and activity symptoms, suggesting that the scale functions differently for the two ethnic groups.

Raju et al. (1995) have suggested that the test characteristic curves from two groups can be utilized to calculate differential test functioning. Differential test functioning can be mathematically illustrated as the following:

$$DTF = \frac{(T_F - T_R)^2}{F}$$

The DTF denotes differential test functioning, E denotes the expectation, T_F denotes the test characteristic of the focal group, and T_R denotes the test characteristic of the reference group (Raju et al., 1995). Holland and Thayer (1988) describe the focal group as the group of primary interest and the reference group as the group in which the focal group will be compared against. In reference to the present study, the following mathematical denotation would apply:

$$DTF = \frac{(T_{FN} - T_{NFN})^2}{F}$$

Whereas, T_{FN} represents the test characteristic curve for the First Nations group and the T_{NFN} represents the test characteristic curve for the non-First Nations group.

It is important to note that differential item functioning does not necessarily indicate the existence of differential test functioning. When examining bias between groups, bias may occur at the item level but not at the scale level. Consequently, the scale may provide an accurate and unbiased measurement for both groups as the bias is eliminated or cancelled out at the scale level (Raju et al., 1995; Shealy & Stout, 1993; Waller et al., 2000). Shealy and Stout (1993) indicate that the cancellation process can occur when the test assessing the latent variable measures dimensions other than what the test is

designed to measure, referred to as the nuisance determinant. Consequently, the nuisance determinant can be multidimensional and some of the dimensions demonstrate bias for one group and other dimensions are biased against the other group. Bias at the test or scale level is depended upon the amount of cancellation that takes place when the two groups are not equivalent on the amount of nuisance determinants. Furthermore, cancellation can eliminate or decrease the amount of bias that is represented at the test or scale level. However, when no cancellation occurs it is possible for some amount of bias at the item level to be amplified at the test or scale level. When differential item functioning amplification occurs, differential test functioning can occur (Shealy & Stout, 1993).

Differential functioning analyses. Item response theory techniques permit different groups of people (e.g., First Nations and non-First Nations) to be compared on items and scales score (Raju et al., 1995; Reise, 1999). In order to examine differential item functioning and differential test functioning, item parameter estimates need to be generated for the two groups, the reference group and the focal group (Flowers et al., 1999). Next, the item parameter estimates from the reference group are linked to a common metric with the item parameters from the focal group using a linear transformation (Flower et al., 1999; Ironson, 1983; Reise, 1999). The calibration of the item parameter estimates by linking the items onto the same metric would be arduous to achieve with traditional statistical analysis techniques (Reise, 1999). Flower et al. (1999) explains that the focal group distribution is used to compute two expected items scores, one score utilizing the item parameter estimates from the focal group and

the other score utilizing the item parameters estimates from the reference group. Consequently, if the two calculated expected scores are not equal, then differential item functioning is evident. The same rational can be applied to differential functioning at the test or scale score level. Expected scale scores are computed for the focal group and one scale score for the reference group. Differential test functioning occurs when there is a difference between the focal group and reference group scale scores (Flowers et al., 1999; Raju et al., 1995).

Scale score differences between the focal and the reference groups are established with significance tests. Consider the following mathematical representation as illustrated by Raju et al. (1995), whereas D represents the normal distribution with a mean of μ_D and a standard deviation of σ_D for respondent s

$$Z_s = \frac{D_s - \mu_D}{\sigma_D}$$

Raju et al. (1995, p. 357) indicate that “it is well known that z^2 has a X^2 distribution with 1 degree of freedom (df), the sum of z^2 across N_F examinees in the focal group has a X^2 distribution with $N_F df$, where N_F is the focal group sample size and N_R is the reference group sample size.” This is represented in the following mathematical representation

$$X^2_{NF} = \sum_{s=1}^{N_F} Z^2 = \frac{\sum_{s=1}^{N_F} (D_s - \mu_D)^2}{\sigma^2_D}$$

The following equation implies that μ_D must be equal to 0

$$\bar{X} (DTF) = \mu_D = 0.$$

Hence, by replacing μ_D the equation is now represented as

$$X^2_{NF} = \frac{\sum_{s=1}^{N_F} D_s^2}{\sigma^2_D}.$$

The mathematical representation for N_F examinees would be denoted as

$$X^2_{NF} = \frac{N_F (DTF)}{\sigma^2_D}$$

The X^2 analysis determines when the differential test functioning statistic is significantly different from 0. When the differential test functioning statistic is below the cut-off value differential test functioning is not evident. Consequently, the Chi-Square analysis is insignificant when the cut-off for the differential test functioning statistic is not exceeded (Flower et al., 1999; Raju et al., 1995). A significant X^2 signifies that one or more items on the test or scale contain bias (Flowers et al., 1999). Raju et al. (1995) proposes that items that contain bias should be removed until X^2 is not significant. Santor and Ramsay (1998) indicate that detecting response differences between groups provides the opportunity to alter test scores to treat individuals impartially and to modify items that may pose as limitations to the assessment scale. Identifying response differences by differential item or test functioning allows for systematic bias to be detected.

Advantages of Item Response Theory

There are many advantages associated with employing item response theory. One of the main advantages of item response theory is the capability of this model to estimate an individual's trait on the same trait scale from any subset of items. The set of items needs to assess a distinct trait, otherwise, if the set of items varies, the trait estimates will have little meaning. As long as the item response model fits the data set, it is possible to obtain an ability estimate for each individual, which will be an unbiased estimate of the individual's true traits (Hambleton & Swaminathan, 1985).

A principal advantage of item response models is the measures are likely to be sample invariant. Sample invariance occurs when the item characteristic curves are not depended upon the sample from which the item data were selected (Hambleton & Swaminathan, 1985; Murphy & Davidshofer, 1998). However, the same cannot be said for classical test theory. An important aspect of item response theory is that the invariance of the measures allows for the characteristics of items to be examined with the characteristics of the individuals responding to the test items without confounding the results, as typically happens in traditional item analyses (Murphy & Davidshofer, 1998).

Another advantage of item response models is the invariance of parameters. This advantage indicates that the comparisons can be made between individuals from different samples and individuals who have completed different subset of items (Hambleton & Swaminathan, 1985). Parameter estimates from different samples (e.g., First Nations and non-First Nations sample) are transformed onto the same scale (Bejar, 1983; Ironson, 1983). Parameter estimates are linearly related to the parameters estimated using another sample

taken from the same population. This occurs despite the distribution ability of the sample upon which the parameters are estimated (Bejar, 1983).

Another advantage of item response model is that it offers a local measure of precision. Item response theory utilizes the information function, in that the model indexes precisely how different the scores are, to characterize the measurement precision. Hence, item response models provide a statistic signifying the precision of each individual's attribute. The trait statistic may differ from one individual to another (Bejar, 1983; Hambleton & Swaminathan, 1985). This is in contrast to classical theory, which relies upon the imprecision of test scores, for instance, the standard error of measurement. Item response models are more beneficial in that it specifically allows certain values of trait or ability to display higher precision (Bejar, 1983). The degree of these advantages depends on the fit between the item response model and the data set. If there is a poor fit between the model and the data, these advantages may be acquired at a lower degree or not at all (Hambleton & Swaminathan, 1985).

Disadvantages of Item Response Theory

Although item response theory has promising psychometric techniques, this procedure is not perfect. Hunter (1975, as cited in Ironson (1983)) indicates that the item characteristic curve represents the observed curve and does not present the true item characteristic curve. There are some faults with utilizing item observed item characteristic curves to identify item bias: (a) larger errors of estimation will be evident for the small numbers of individuals with very high or very low attributes who are represented at the high and low portions of the curves, and (b) item characteristic curves for an item may not be a reliable

estimate and will appear to have a poorer discrimination than in actuality. Raju et al. (1995) also indicates that error may also occur in differential test functioning from the use of person and item parameter estimates, and from sampling error.

Purpose of the Present Study

The available research literature poses important questions regarding the applicability of utilizing conventional instruments of psychopathology with First Nations clients (Allen, 1998). It is difficult to conclude if elevated symptomatology is an indication of psychopathology or cultural differences (Dillard & Manson, 2000), produced by differential item or test functioning. An inaccurate assessment of psychopathology may result in serious consequences such as misdiagnosis and inappropriate treatment interventions for First Nations clients. It is important for mental health practitioners to base their assessment practices on research evidence (Hunsley, 2003; Lilienfeld et al., 2003); however there is no scientific evidence regarding the applicability of utilizing the BCFPI with First Nations children and adolescents. Hence, the purpose of the present analyses were to examine cultural measurement bias in the performance of the BCFPI for First Nations and non-First Nations children and adolescents using IRT models.

The present study utilized item response theory techniques to assess for differential item functioning and differential test functioning between First Nations and non-First Nations children and adolescents on the mental health scales of the BCFPI. This provided the opportunity to examine whether the items and total scale scores on the BCFPI function similarly for First Nations and non-First Nations children and adolescents. These sophisticated psychometric

techniques can determine whether equivalent total scale scores for First Nations and non-First Nations children and adolescents have the same indication of psychopathology for the two groups. Item response techniques allows for First Nations and non-First Nations children and adolescents to be compared on the same latent traits, consequently indicating whether culture influences the item and test responses on the BCFPI for First Nations children and adolescents. In addition, the results will indicate whether the BCFPI can assess emotional and behavioural symptoms in an accurate and unbiased manner among First Nations children and adolescents. Establishing cultural differences between First Nations and non-First Nations children and adolescents will provide the opportunity to eliminate systematic biases in the BCFPI by improving specific items and correcting scores to attend to the limitations on each of the scales. Greater understanding of the response styles will assist in accurately identifying assessment problems within First Nations children and adolescents. This clinical information is imperative if we are to respond effectively to the mental health needs of First Nations children and adolescents within our health care system.

Although no previous research exists that has utilized differential item and test functioning to assess the utility of the BCFPI among the First Nations population, it was hypothesized that First Nations and non-First Nations children and adolescents would respond differentially to the items and scales of the BCFPI. Although, there is also a lack of research that has examined First Nations children's mental health, the hypothesis is based upon research among the adult First Nations population. Scale scores differences (using classical test theory) have indicated that adult First Nations individuals score higher than their non-

First Nations counterparts on measures of depression, hypomania, hypochondriasis, hysteria, schizophrenia, paranoia, psychopathic deviate, psychasthenia, and social introversion. Moreover, researchers have speculated that these differences may be influenced by cultural variations in test response (Borzecki et al., 1984; Butcher et al., 1983; Graham, 1993; Herreid & Herreid, 1966; Hoffman et al., 1985; Kline et al., 1973; Mandelzys & Lane, 1980; Page & Bozlee, 1982; Panton, 1980; Pollack & Shore, 1980; Uecker et al., 1980; Venn, 1988). Consequently, the purpose of the present study was to determine whether the commonly observed higher psychopathology scores for First Nations individuals may be due to bias at the item or scale levels and to examine whether the group mean differences remained or disappeared after correcting for any differential item or test functioning. In other words, are the commonly reported higher elevations due to test bias? Or do group differences in elevations still exist after corrections for test bias? It was hypothesized that differences would be noted at the item and test levels of the BCFPI between First Nations and non-First Nations children and adolescents, as adults' measures of psychopathology have indicated group differences between First Nations and non-First Nations cultural groups.

Method

Participants

There were a total of 3,265 BCFPI parent reports included in the study. There were 340 First Nations BCFPI parent reports. The First Nations data were obtained from the BCFPI database at Dilico Ojibway Child and Family services, an agency that offers wholistic services in Mental Health, Addictions, Health, and Child Welfare to Anishinabek children, families, and communities in Northern Ontario. The names of the clients were not exported from the BCFPI database to

ensure anonymity and confidentiality. The non-First Nations BCFPI data were obtained from Dr. Michael Boyle and Yvonne Racine, the original authors of the Revised Ontario Health Study (Boyle et al., 1993a), in which the BCFPI clinical and population norms are based upon. There were 1242 parent reports included from the clinical norm sample and 1683 parent reports from the general population norm sample. The average age of the participants was 10.96 years ($SD = 3.59$).

Measure

Brief Child and Family Phone Interview (BCFPI; Cunningham, Pettingill, & Boyle, 2004). As mentioned previously, the BCFPI is a structured interview that is administered to parents, teachers, or adolescents to assess emotional and behavioural disorders in children and adolescents. The structured interview is based upon diagnostic criteria and evaluates regulation of attention, impulsiveness, and activity level, cooperation with others, conduct, separation from parents, managing anxiety, and managing mood. The respondents rated the symptoms of psychopathology on the following scale: 0 = "never or not true," 1 = "sometimes true," or 2 = "often or very true."

The BCFPI has population and clinical norms that are based upon the general population and clinical samples from the Revised Ontario Child Health Study (Boyle et al., 1993; Cunningham et al., 2001). The results of the BCFPI are presented as *t*-scores that have a mean of 50 and a standard deviation of 10. Children's scores are compared to the norms of children aged 6 to 12 years or 13 to 18 years of age. Children between the ages of 3 to 5 years are compared to the 6

to 12 years old children, consequently these results must be cautiously interpreted.

As mentioned previously, the items on the BCFPI were selected from a larger pool of items from the Revised Ontario Child Health Study. The scales for the BCFPI were statistically derived using principal components factor analyses with varimax rotations on the general population sample and the clinical sample from the Revised Ontario Child Health Study (Boyle et al., 1993a, 1993b; Cunningham et al., 2001). Table 1 illustrates the BCFPI questions for each scale. The factor loadings for the BCFPI Regulating Attention, Impulsivity, and Activity Level scale ranged from .606 to .758 for the general population sample of the Revised Ontario Child Health Study items that assessed attention-deficit hyperactivity disorder and .577 to .804 for the clinical sample. The internal consistency for this scale with the Revised Ontario Child Health Study population sample was .83 and .84 for the clinical sample. The content validity of this scale was established by selecting questions that query inattentive and hyperactive-impulsive ADHD symptoms as illustrated in the DSM-IV. As the BCFPI scales are abbreviated from the Revised Ontario Child Health Study scales, correlations were conducted between the scales to establish concurrent validity for the BCFPI. The concurrent validity coefficient for the BCFPI Regulating Attention, Impulsivity, and Activity Level scale was .91 for both the Revised Ontario Child Health Study population and clinical samples (Boyle et al., 1993; Cunningham et al., 2001).

The Cooperation with Others scale was statistically derived using principal components factor analyses with varimax rotations on the general population

sample and the clinical sample from the Revised Ontario Child Health Study (Boyle et al., 1993a, 1993b; Cunningham et al., 2001). The factor loadings for the Cooperation with Others scale ranged from .474 to .745 for the Revised Ontario Child Health Study items that assessed oppositional disorder in the population sample, and from .644 to .749 for the clinical sample. The cronbach's alpha score for this scale with the Revised Ontario Child Health Study population sample was .82 and .85 for the clinical sample. The content validity of this scale was generated by choosing items that depict oppositional defiant DSM-IV symptoms. The concurrent validity of the Cooperation with Others scale was established with the items from the Revised Ontario Child Health Study and the coefficients was .95 for both the population and clinical samples (Boyle et al., 1993; Cunningham et al., 2001).

The Conduct scale was statistically derived using principal components factor analyses with varimax rotations on the general population sample and the clinical sample from the Revised Ontario Child Health Study (Boyle et al., 1993a, 1993b; Cunningham et al., 2001). The factor loadings for the Conduct scale ranged from .364 to .847 for the Revised Ontario Child Health Study items that assessed conduct disorder in the population sample and .583 to .631 for the clinical sample. The internal consistency for the Conduct scale in the Revised Ontario Child Health study population sample was .56 and .73 for the clinical sample. The content validity of this scale was established by selecting questions that query conduct disorder symptoms as illustrated in the DSM-IV. The concurrent validity of the Conduct scale was established with items from the Revised Ontario Child Health Study, the coefficients were .81 for the population

sample and .89 for the clinical sample (Boyle et al., 1993; Cunningham et al., 2001).

The Separation from Parents scale was statistically derived using principal components factor analyses with varimax rotations on the general population sample and the clinical sample from the Revised Ontario Child Health Study (Boyle et al., 1993a, 1993b; Cunningham et al., 2001). The factor loadings for the Separation from Parents scale ranged from .457 to .760 for the Revised Ontario Child Health Study items that assessed separation anxiety disorder in the population sample and from .609 to .794 for the clinical sample. The cronbach's alpha scores in the Revised Ontario Child Health Study population sample for the Separation from Parents scale was .75 and .81 for the clinical sample. The content validity of this scale was generated by choosing items that depict separation anxiety symptoms as described in the DSM-IV. The concurrent validity coefficient for the Separation from Parents scale was .95 for the Revised Ontario Child Health Study population sample and .96 for the clinical sample (Boyle et al., 1993; Cunningham et al., 2001).

The Managing Anxiety scale was statistically derived using principal components factor analyses with varimax rotations on the general population sample and the clinical sample from the Revised Ontario Child Health Study (Boyle et al., 1993a, 1993b; Cunningham et al., 2001). The factor loadings for the Managing Anxiety scale ranged from .476 to .739 for the Revised Ontario Child Health Study items that assessed overanxious disorder in the population sample and from .636 to .782 for the clinical sample. The internal consistency score in the Revised Ontario Child Health Study population sample for the Managing

Anxiety scale was .77 and .82 for the clinical sample. The content validity of this scale was established by selecting questions that query anxiety disorder symptoms as illustrated in the DSM-IV. The concurrent validity of the Managing Anxiety scale was established with items from the Revised Ontario Child Health Study, the coefficients were .83 for the population sample and .84 for the clinical sample (Boyle et al., 1993; Cunningham et al., 2001).

The Managing Mood scale was statistically derived using principal components factor analyses with varimax rotations on the general population sample and the clinical sample from the Revised Ontario Child Health Study (Boyle et al., 1993a, 1993b; Cunningham et al., 2001). The factor loadings for the Managing Mood scale ranged from .343 to .745 for the Revised Ontario Child Health Study items that assessed depressive symptoms in the population sample and .623 to .724 for the clinical sample. The cronbach's alpha score for the Managing Mood scale in the Revised Ontario Child Health Study population sample was .78 and .85 for the clinical sample. The content validity of this scale was generated by choosing items that depict major depressive symptoms as depicted in the DSM-IV. The concurrent validity of the Managing Mood scale was established with items from the Revised Ontario Child Health Study, the coefficients were .77 for the population sample and .82 for the clinical sample (Boyle et al., 1993; Cunningham et al., 2001).

The internal consistency score for the Child Functional Impact scale in the Revised Ontario Child Health Study population sample was .86 and .74 for the Revised Ontario Child Health Study clinical sample. The cronbach's alpha score for the Family Functional Impact scale in the Revised Ontario Child Health Study

population sample was .69 and .78 for the Revised Ontario Child Health Study clinical sample (Boyle et al., 1993; Cunningham et al., 2001).

The present analyses focused on the parent reports of the BCFPI that were completed at intake; consequently the teacher and youth reports were not included in the analyses. The data from the scales that assess emotional and behavioural symptoms among First Nations and non-First Nations children and adolescents were assessed for differential item functioning and differential test functioning using IRT analyses.

Results

Sample Characteristics

It is important to note that, as Dilico Ojibway Child and Family Services has a child protection department, 37 of the participants from this sample were in foster care. It is unknown how many children from the comparison sample were in foster care. Table 2 reports the demographic information for the First Nations clinical sample, non-First Nations clinical sample, and the non-First Nations general population sample. The demographic information included whether the respondent was a single parent, the language most often spoken in the home, highest level of education for the respondent, highest level of education for the respondent's spouse, and the total family income before taxes. The results from the demographic information indicate that a higher percentage of the First Nations respondents consisted mostly of double parent households, whereas the non-First Nations clinical and community samples were mostly from a single parent dwelling. The language most often spoken in all three samples was English. Most of the First Nations sample and the non-First Nations clinical and

community samples had some secondary education. The most frequently endorsed total family income before taxes for First Nations individuals was less than \$10,000 (15.9%), while the non-First Nations clinical sample was more than \$60,000 (17.6%), and the non-First Nations general population sample was between \$40,000 and \$49,000 (19.3%).

Group Differences In Raw Score Means

Independent *t*-tests were conducted and effect sizes were calculated to examine total scale score differences between the First Nations clinical sample and non-First Nations clinical and general population samples on the scales of the BCFPI. Rosnow and Rosenthal (1988) indicate that effect sizes should be calculated with *t* tests in order to increase statistical power and decrease the probability of making a Type II error. Consequently, the following equation was utilized to determine effect sizes:

$$r = \sqrt{t^2 / (t^2 + df)}$$

The t^2 represents the significance *t*-test and *df* represents degrees of freedom. Cohen (1977, as cited in Rosnow & Rosenthal, 1988) indicates that effect sizes of .1 can be interpreted as "small," .3 can be interpreted as "medium," and .5 can be interpreted as "large."

Table 3 lists the means, standard deviations, range of values, *t* statistics, and effect sizes for the First Nations clinical sample and non-First Nations clinical sample on the measures of Regulation of Attention, Impulsivity, and Activity Level, Cooperativeness, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood. The First Nations clinical sample and non-First Nations clinical sample total scale scores were not statistically significant for the

Regulation of Attention, Impulsivity, and Activity Level scale, $t(1522) = -.02, p = .984$, and an effect size of $r = .002$. Statistically significant differences were found between First Nations clinical sample and non-First Nations clinical sample on the Cooperativeness with Others scale, $t(510.777) = -2.24, p = .025$, and an effect size of $r = .099$; Conduct scale, $t(528.939) = 2.04, p = .042$, and an effect size of $r = .088$; Separation from Parents scale, $t(565.320) = 2.16, p = .032$, and an effect size of $r = .090$; Managing Anxiety scale, $t(546.374) = -4.93, p = .000$, and an effect size of $r = .206$; and Managing Mood scale, $t(552.948) = -8.17, p = .000$, and an effect size of $r = .328$.

Table 4 lists the means, standard deviations, range of values, t statistics, and effect sizes for the measures of Regulation of Attention, Impulsivity, and Activity Level, Cooperativeness, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood between the First Nations clinical sample and the non-First Nations general population sample. The First Nations clinical sample reported higher means on all scales of the BCFPI when compared to the non-First Nations general population sample. Statistically significant differences were found between First Nations clinical sample and non-First Nations general population sample on the Regulating Attention, Impulsivity, and Activity Level scale, $t(423.839) = 16.16, p = .000$, and an effect size of $r = .617$; Cooperativeness with Others scale, $t(408.221) = 17.25, p = .000$, and an effect size of $r = .649$; Conduct scale, $t(350.303) = 14.54, p = .000$, and an effect size of $r = .613$; Separation from Parents scales, $t(402.343) = 12.24, p = .000$, and an effect size of $r = .521$; Managing Anxiety scale, $t(401.437) = 3.36, p = .001$, and an effect

size of $r = .165$; and Managing Mood scale, $t(362.974) = 12.30$, $p = .000$, and an effect size of $r = .542$.

Table 5 lists the means, standard deviations, range of values, t statistics, and effect sizes for the measures of Regulation of Attention, Impulsivity, and Activity Level, Cooperativeness, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood between the non-First Nations clinical sample and the non-First Nations general population sample. The non-First Nations clinical sample reported higher means on all scales of the BCFPI when compared to the non-First Nations general population sample. Statistically significant differences were found between non-First Nations clinical sample and non-First Nations general population on the Regulating Attention, Impulsivity, and Activity Level scale, $t(2252.68) = 27.63$, $p = .000$, and an effect size of $r = .503$; Cooperativeness with Others scale, $t(2172.60) = 34.24$, $p = .000$, and an effect size of $r = .592$; Conduct scale, $t(1423.41) = 22.34$, $p = .000$, and an effect size of $r = .510$; Separation from Parents scale, $t(1907.50) = 15.52$, $p = .000$, and an effect size of $r = .335$; Managing Anxiety scale, $t(1970.04) = 13.90$, $p = .000$, and an effect size of $r = .299$; and Managing Mood scale, $t(1616.98) = 34.68$, $p = .000$, and an effect size of $r = .653$.

Item Response Theory Nonparametric Analyses

Item inter-item correlations were conducted to ensure the assumption of scale homogeneity was met. Cronbach's alphas were calculated to assess the internal consistency of the mental health scales on the BCFPI. The coefficient alphas were as follows for the mental health scales of the BCFPI: Regulation of Attention, Impulsivity, and Activity scale, $\alpha = .8671$; Cooperativeness with Others

scale, $\alpha = .8805$, Conduct scale, $\alpha = .7442$; Separation from Parents scale, $\alpha = .8059$; Managing Anxiety scale, $\alpha = .8106$; and Managing Mood scale, $\alpha = .8628$. All reliability coefficients are sufficiently high and suggest that the assumption of scale homogeneity was met.

Nonparametric kernel-smoothing techniques were utilized to assess for differential item functioning or item bias on the BCFPI among First Nations and non-First Nations children and adolescents. TESTGRAF software (Ramsay, 1993) was used to generate item characteristic curves to evaluate differential item functioning and differential test functioning. Figure 7 illustrates item characteristic curves for the items on the BCFPI scales Regulation of Attention, Impulsivity, and Activity Level among the First Nations clinical sample and the non-First Nations clinical sample. Figure 8 depicts item characteristic curves for the items on the BCFPI scale Cooperativeness with Others among the First Nations clinical sample and the non-First Nations clinical sample. Figure 9 portrays the item characteristic curves for the Conduct scale of the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample. Figure 10 portrays the item characteristic curves for the Separation from Parents scale among the First Nations clinical sample and the non-First Nations clinical sample. Figure 11 illustrates the item characteristic curves for the BCFPI scale Managing Anxiety among the First Nations clinical sample and the non-First Nations clinical sample. Figure 12 depicts item characteristic curves for the Managing Mood scale among the First Nations clinical sample and the non-First Nations clinical sample. Nonparametric kernel-smoothing techniques were also employed to examine item bias on the BCFPI among the First Nations clinical

sample and non-First Nations general population sample. Figure 13 illustrates item characteristic curves for the items on the BCFPI scale Regulation of Attention, Impulsivity, and Activity among the First Nations clinical sample and the non-First Nations community sample. Figure 14 depicts item characteristic curves for the items on the BCFPI subscale Cooperativeness with Others among the First Nations clinical sample and the non-First Nations community sample. Figure 15 portrays the item characteristic curves for the Conduct scale of the BCFPI among the First Nations clinical sample and the non-First Nations community sample. Figure 16 portrays the item characteristic curves for the Separation from Parents scale among the First Nations clinical sample and the non-First Nations community sample. Figure 17 illustrates the item characteristic curves for the BCFPI scale Managing Anxiety among the First Nations clinical sample and the non-First Nations community sample. Figure 18 depicts item characteristic curves for the Managing Mood scale among the First Nations clinical sample and the non-First Nations community sample.

Item Response Theory Parametric Analyses

In addition to assessing differential item functioning and differential test functioning with nonparametric analyses, parametric techniques were also utilized to quantify differential item functioning and differential test functioning between First Nations and non-Nations children and adolescents on the mental health scales of the BCFPI.

The first step of this analyses consisted of utilizing the computer program MULTILOG 7.0 written by David Thissen to estimate item parameters. This software permits item analyses in the context of item response models, as it is

possible to examine items with multiple response options and utilize logistic response models. This program can examine data sets with item response options that are dichotomous or polytomous. MULTILOG can generate estimate parameters of 1, 2, or 3 parameter logistic models, Samejima's graded response model, Bock's nominal response model, or Master's partial credit model. For these analyses the graded response model was utilized (du Toit, 2003; Embretson & Reise, 2000). Samejima's graded response model presumes an ordered response, in which the respondent is able to only select one response value per item and the item slope and trait parameters are considered in the analyses (Flowers et al., 1999). MULTILOG provided marginal maximum likelihood item parameter estimates for the items on the mental health scales of the BCFPI. The item parameter estimates were performed separately for the First Nations clinical sample, non-First Nations clinical sample, and non-First Nations general population sample (du Toit, 2003; Embretson & Reise, 2000).

The next step of the analyses consisted of establishing a common metric for the item parameters. The computer program EQUATE (Baker, 1993) used the graded response model to link or calibrate the item parameters with a modified test characteristic curve method. The EQUATE software incorporates the item parameters estimates (generated from MULTILOG) within the "from" or reference group to the metric of the "to" or focal group. Consequently, the item parameter estimates of the First Nations clinical sample were equated to the metric of the non-First Nations clinical sample to calculate the slope and intercept coefficients of the linear transformation of each of the mental health scales on the BCFPI. The item parameter estimates of the First Nations clinical

sample were also transformed to the metric of the non-First Nations general population sample to compute the slope and intercept coefficients of the linear transformation of the BCFPI mental health scale (Baker, 1993).

The last step within the IRT analyses consisted of utilizing the computer program DFIT written by Nambury Raju (1995). This program computes the differential item functioning and differential test functioning statistics. Table 6 displays the differential item functioning statistics between the First Nations clinical sample and the non-First Nations clinical sample on the Regulation of Attention, Impulsivity, and Activity Level, Cooperativeness with Others, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood scales on the BCFPI. The differential item functioning analyses indicated that all six items on the Regulation of Attention, Impulsivity, and Activity Level scale contained bias. The differential item functioning analyses of the Cooperativeness with Others scale indicated that five of the six items contain bias. The differential item functioning analyses of the Conduct scale indicated that four of the six items contain bias. The differential item functioning analyses of the Separation from Parents scale indicated that all six items from the scale contain bias. The differential item functioning analyses of the Managing Anxiety scale indicated that all six items contain bias. The differential item functioning analyses of the Managing Mood scale indicated that eight out of the nine items contain bias. The results suggest that most items on all mental health scales of the BCFPI contain bias.

Table 7 displays the differential item functioning statistics between the First Nations clinical sample and the non-First Nations general population

sample on the Regulation of Attention, Impulsivity, and Activity Level, Cooperativeness with Others, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood scales on the BCFPI. The differential item functioning analyses indicated that all six items of the Regulation of Attention, Impulsivity, and Activity Level scale contained bias. The differential item functioning analyses of the Cooperativeness with Other scale indicated that all six items from the scale contain bias. The differential item functioning analyses of the Conduct scale indicated that five out of the six items contained bias. The differential item functioning analyses of the Separation from Parents scale indicated that five out of six items contained bias. The differential item functioning of the Managing Anxiety scale indicated that all six items on the scale contained bias. The differential item functioning analyses of the Managing Mood scale indicated that eight of the nine items contain bias. Overall the results suggest that most items on the mental health scales of the BCFPI contain bias.

Table 8 displays the number of items on each scale that contain bias and the differential test functioning statistics between the First Nations clinical sample and the non-First Nations clinical sample on the Regulation of Attention, Impulsivity, and Activity Level, Cooperativeness with Others, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood scales on the BCFPI. The differential item functioning analyses indicated that all six items on the Regulation of Attention, Impulsivity, and Activity Level scale contained bias. However, the differential test functioning analyses of the same scale indicated that the differential test functioning statistic of .05338 did not surpass the cut-off value of .144. The differential item functioning analyses of the Cooperativeness

with Others scale indicated that five of the six items contain bias. However, the differential test functioning analyses of the Cooperativeness with Others scale indicated that the differential test functioning statistic of .00866 did not exceed the cut-off value of .144. The differential item functioning analyses of the Conduct scale indicated that four of the six items contain bias. However, the differential test functioning analyses of the Conduct scale indicated that the differential test functioning statistic of .03145 did not surpass the cut-off value of .144. The differential item functioning analyses of the Separation from Parents scale indicated that all six items from the scale contain bias. However, the differential test functioning analyses of the Separation from Parents scale indicated that the differential test functioning statistic of .00137 did not exceed the cut-off value of .144. The differential item functioning analyses of the Managing Anxiety scale indicated that all six items contain bias. However, the differential test functioning analyses of the Managing Anxiety scale indicated that the differential test functioning statistic of .00404 did not surpass the .144 cut-off value. The differential item functioning analyses of the Managing Mood scale indicated that eight out of the nine items contain bias. However, the differential test functioning analyses of the Managing Mood scale indicated that the differential test functioning statistic of .02088 did not exceed the cut-off value of .216. These findings indicated that although most items exhibited differential item functioning the bias is not sustained at the test or scale level. These results suggest that there are no significant scale score differences between First Nations and non-First Nations children and adolescents on the mental health scales of the BCFPI.

Table 9 displays the number of items that contain differential item functioning and the differential test functioning statistic for First Nations clinical sample and the non-First Nations general population sample on the Regulation of Attention, Impulsivity, and Activity Level, Cooperativeness with Others, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood scales on the BCFPI. The differential item functioning analyses indicated that all six items of the Regulation of Attention, Impulsivity, and Activity Level scale contained bias. However, the differential test functioning analyses of the Regulation of Attention, Impulsivity, and Activity Level scale indicated that the differential test functioning statistic of .01608 did not surpass the cut-off value of .144. The differential item functioning analyses of the Cooperativeness with Other scale indicated that all six items from the scale contain bias. However, the differential test functioning analyses of the Cooperativeness with Others scale indicated that the differential test functioning statistic of .05511 did not exceed the cut-off value of .144. The differential item functioning analyses of the Conduct scale indicated that five out of the six items contained bias. However, the differential test functioning analyses of the Conduct scale indicated that the differential test functioning statistic of .00002 did not surpass the cut-off value of .144. The differential item functioning analyses of the Separation from Parents scale indicated that five out of six items contained bias. However, the differential test functioning analyses of the Separation from Parents scale indicated that the differential test functioning statistic of .01561 did not exceed the cut-off value of .144. The differential item functioning of the Managing Anxiety scale indicated that all six items on the scale contained bias. However, the differential test

functioning analyses of the Managing Anxiety scale indicated that the differential test functioning statistic of .02463 did not surpass the .144 cut-off value. The differential item functioning analyses of the Managing Mood scale indicated that eight of the nine items contain bias. However, the differential test functioning analyses of the Managing Mood scale indicated that the differential test functioning statistic of .00301 did not exceed the cut-off values of .216. Overall the results suggest that most items on the mental health scales contain bias. However, this bias is not sustained at the test or scale level indicating that there are no significant test or scale score differences between First Nations and non-First Nations children and adolescents on the mental health scales of the BCFPI.

Discussion

The present analyses address important clinical issues of measurement invariance in the BCFPI among First Nations children and adolescents. To the authors' knowledge, this is the first application of item response techniques to identify cultural differences in the assessment of emotional and behavioural symptoms with the BCFPI. It is surprising that no other researchers have assessed measurement bias in the BCFPI, as this measure is mandated by the government to be completed by all mental health providers for children and adolescents in the province of Ontario. The purpose of the present study was to examine measurement item and scale bias on the mental health scales of the BCFPI among First Nations and non-First Nations children and adolescents. This analyses utilized item response theory methods to conduct differential item and test functioning among the Regulation of Attention, Impulsivity, and Activity Level, Cooperativeness with Others, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood scales on

the BCFPI. Item response theory analyses permits First Nations and non-First Nations children and adolescents to be equally compared on the scales of the BCFPI.

Mean Score Differences

The results of the scale score analyses between First Nations and non-First Nations children and adolescents from clinical samples indicated significant differences between the groups on most scales of the BCFPI. Independent *t*-test analyses indicated scale score differences on the Cooperativeness with Others, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood scales of the BCFPI. The First Nations respondents reported higher group mean scores on the Conduct scale and Separation from Parents scale. The non-First Nations respondents reported higher mean scores on the Cooperativeness with Others, Managing Anxiety, and Managing Mood scales. Differences between the groups were not evident on the Regulation of Attention, Impulsivity, and Activity Level scales. However, the comparisons of scale scores differences does not indicate true scale score differences between the groups. Classical test theory does not differentiate mean score differences from measurement bias. Item response theory can make this distinction by comparing the groups on latent variables by establishing a common metric for the item parameters (Kim, Pilkonis, Frank, Thase, & Reynolds, 2002). Consequently, the group differences attained with classical test theory were not sustained with the item response theory analyses, as there was no evidence of differential test functioning or bias at the total scale level. The item response theory analyses will be discussed in the next section.

The results of the independent *t*-test analyses between the First Nations clinical sample and the non-First Nations general population sample reveals scale score differences on the Regulating Attention, Impulsivity, and Activity Level, Cooperativeness with Others, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood scales on the BCFPI. The First Nations clinical sample reported higher means on all scales of the BCFPI when compared to the non-First Nations general population sample. These results are to be expected as the analyses is comparing respondents from mental health agencies to respondents from the general community.

Independent *t*-test analyses were also conducted for non-First Nations clinical sample and the non-First Nation general population sample on the mental health scales of the BCFPI. The results reveal scale score differences on the Regulating Attention, Impulsivity, and Activity Level, Cooperativeness with Others, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood scales on the BCFPI. When compared to the non-First Nations general population sample, the non-First Nations clinical sample reported significantly higher means. Again, these results are not surprising given the analyses compared a clinical sample and a community sample.

Differential Functioning Analyses

The present study addresses an important clinical issue of whether the BCFPI can assess emotional and behavioural disorders among First Nations children and adolescents in an unbiased manner. Lord and Novick (1968) explained that item response models theorize that responses to assessments are explained by an individual's unobservable or latent trait, therefore, representing

the observable item as measured by the assessment. Consequently, one would theorize that the BCFPI assesses the following latent traits: regulating attention, impulsivity, and activity level, cooperating with other people, conduct, separation from parents, anxiety, and depressive symptoms. The endorsements of the latent traits were expressed in item characteristic curves that represent the relationship between the respondent's latent traits and the endorsement of the item.

Consequently, respondents with elevated traits were likely to endorse items that reflect those traits (Hambleton & Swaminathan, 1985; Murphy & Davidshofer, 1998; Reise, 1999). Those who possessed more of the latent traits as assessed by each of the six mental health scales on the BCFPI endorsed more of the items on each scale.

The differential item functioning analyses indicated that significant bias does exist for most items on all mental health scales of the BCFPI between First Nations and non-First Nations children and adolescents. However, the differential test functioning analyses has indicated that these item differences are not sustained at the scale level. The findings from the differential item and test functioning will be discussed for each scale.

Regulating Attention, Impulsivity, and Activity Level scale. Differential item functioning was assessed on the six-item Regulation Attention, Impulsivity, and Activity Level scale between First Nations and non-First Nations children and adolescents from the clinical samples. The results indicated that the following questions exhibited item bias: "Distractible, has trouble sticking to an activity," "Fails to finish things he starts," "Has difficulty following directions or instructions," "Impulsive, acts without stopping to think," "Jumps from one

activity to another,” and “Fidgets.” This suggests that these items assessing attention, impulsivity, and activity level from the BCFPI function differently for First Nations and non-First Nations children and adolescents. However, the differential test functioning statistic did not exceed the cut-off value required for there to be any meaningful bias at the scale level. Therefore, bias is not present at the scale level of the Regulation of Attention, Impulsivity, and Activity Level scale of the BCFPI.

Cooperativeness with Others scale. Differential item functioning analyses were conducted on the six-items from the Cooperativeness with Others scale between First Nations and non-First Nations children and adolescents from the clinical samples. The analyses indicated that item bias was evident on the following questions: “Cranky,” “Defiant, talks back to adults,” “Blames other for own mistakes,” “Easily annoyed by others,” and “Agues a lot with adults.” These results indicate that the items assessing children and adolescent’s ability to cooperate with others on the BCFPI functions differently for First Nations and non-First Nations children and adolescents. However, the item “Angry and resentful” from the Cooperativeness with Others scale, did not exhibit any bias between the First Nations and non-First Nations children and adolescents. Although, bias was significant at the item level, the differential test functioning analyses suggested bias was not significant at the scale level as the differential test functioning statistic did not surpass the cut-off value. Scale level bias was not evident on the Cooperativeness with Others scale between First Nations and non-First Nations children and adolescents.

Conduct scale. Item differences were assessed with differential item functioning between First Nations and non-First Nations children and adolescents from the clinical samples on the six-items from the Conduct scale of the BCFPI. The differential item functioning analyses indicated item bias on the following questions from the Conduct scale: “Steals things at home,” “Destroys things belonging to others,” “Engages in vandalism,” and “Has broken into a house, building or car.” This suggests that these items that assess conduct behaviour on the BCFPI function differently for First Nations and non-First Nations children and adolescents. The questions “Physically attacks people” or “Uses weapons when fighting” did not exhibit any bias between First Nations and non-First Nations children and adolescents. This suggests that these items are an accurate assessment of these symptoms among First Nations children and adolescents. However, the differential test functioning analyses indicated that the differential test functioning statistic did not exceed the cut-off value required for differential test functioning. This result suggests that bias at the scale level of the Conduct scale is not evident. The scale scores between First Nations and non-First Nations children and adolescents do not significantly differ.

Separation from Parents scale. Differential item functioning analyses were conducted on six-items from the Separation from Parents scale between First Nations and non-First Nations children and adolescents from the clinical samples. The results indicated that the following questions exhibited item bias: “Worries bad things will happen to loved ones,” “Worries about being separated from loved ones,” “Scared to sleep without parents nearby,” “Overly upset when leaving loved ones,” “Overly upset while away from loved ones,” and “Complains

of feeling sick before separating.” These results suggest that these items assessing the child or adolescent’s comfort with separating from their parents function differently for First Nations and non-First Nations children and adolescents. However, the differential test functioning analyses indicated that the differential test functioning statistic did not surpass the required cut-off value for differential test functioning to be evident. As a result, there were no significant differences or bias between First Nations and non-First Nations children and adolescents on the Separation from Parents scale.

Managing Anxiety scale. Item differences between First Nations and non-First Nations children and adolescents on the six-items from the Managing Anxiety scale were assessed with differential item functioning. The analyses indicated item bias was evident in the following questions: “Worries about doing better at things,” “Worries about past behaviour,” “Worries about doing the wrong thing,” “Worries about things in the future,” “Is afraid of making mistakes,” and “Is overly anxious to please people.” These results indicate that the assessment of anxiety symptoms from the BCFPI do not function similarly for First Nations and non-First Nations children and adolescents. However, the differential test functioning analyses indicated that the differential test functioning statistic did not exceed the cut-off value required for differential test functioning to be evident. Subsequently, the total scale score value for the Managing Anxiety scale did not significantly differ between First Nations and non-First Nations children and adolescents.

Managing Mood scale. Differential item functioning analyses were conducted on the nine-items of the Managing Mood scale between First Nations

and non-First Nations children and adolescents from the clinical samples. The results from the differential item functioning analyses indicated that the following questions exhibited item bias: “No interest in usual activities,” “Gets no pleasure from usual activities,” “Has trouble enjoying self,” “Not as happy as other children,” “Feels hopeless,” “Unhappy, sad or depressed,” “Has lost a lot of weight without trying,” and “Deliberately harms self or attempts suicide.” This result suggests that the assessment of depressive symptoms from the BCFPI function differently for First Nations and non-First Nations children and adolescents. There was no item bias for the item “Talks about killing himself/herself.” This result indicates that this assessment of suicidal ideation functions similarly for First Nations and non First Nations children and adolescents. Although bias was found at the item level, bias was not evident at the total scale level, as the differential test functioning statistic did not surpass the cut-off value.

Differential item functioning was found for most items on all mental health scales of the BCFPI for First Nations and non-First Nations children and adolescents. Although bias was found for most items on the BCFPI mental health scales, it is important to note that most clinical practice is not based upon item response. Rather, clinical practice is based upon total scale scores; consequently for clinical purposes it is important to examine the aggregated items as total scale scores. The differential test functioning analyses for all scales indicated that the biased items were cancelled out or amplified at the scale level (Waller et al., 2000). When the items are summed together to generate the test characteristic, the bias from the items is eliminated or minimized when examining the scale

level differences (Raju et al., 1995). Consequently, “the presence of differential item functioning does not lead inexorably to differential test functioning” (Waller et al., 2000, p. 143,). Hence, the analyses of the BCFPI mental health scales indicate that the First Nations and non-First Nations children and adolescents can be compared with the assurance that the differences indicated are a result of true group differences and not due to measurement bias.

Culture Differences and the BCFPI

Researchers have indicated that conventional assessment measures that utilize standardized norms are likely to result in invalid comparisons in minority populations (Bonder et al., 2002). This does not appear to be the case with the BCFPI application among First Nations children and adolescents. Differential test functioning analyses indicated that First Nations and non-First Nations children and adolescents can be accurately assessed at the scale level when compared to non-First Nations children and adolescents.

There is no available child and adolescent research to compare these present research findings. However, researchers who have utilized classical test theory, have found that adult First Nations participants reported substantially elevated scores on depression, hypomania, hypochondriasis, hysteria, schizophrenia, paranoia, psychopathic deviate, psychasthenia, and social introversion when compared to their dominant culture counterpart on instruments of psychopathology. Researchers have contested that these psychological profiles are substantially influenced by cultural variations in test response (Borzecki et al., 1984; Butcher et al., 1983; Graham, 1993; Herreid & Herreid, 1966; Hoffman et al., 1985; Kline et al., 1973; Mandelzys & Lane, 1980; Page & Bozlee, 1982; Pantan, 1980; Pollack & Shore, 1980; Uecker et al., 1980;

Venn, 1988). With the application of item response techniques, First Nations children and adolescents reported compared scale scores on the Regulation of Attention, Impulsivity, and Activity, Cooperativeness with Others, Conduct, Separation from Parents, Managing Anxiety, and Managing Mood scales of the BCFPI.

One could speculate that the item response variations between First Nations and non-First Nations children and adolescents could be attributed to differences in culture. Tseng (2001) acknowledges that culture significantly impacts how mental health symptoms are communicated or presented to the clinician. An important component of communicating is language and language may play an important role in evaluating item response differences. Kirmayer et al. (2000) indicates that “language is a basic conveyor of culture and most people are connected to their emotions and intimate thoughts most readily in their first language or language of everyday life” (p. 17). The language spoken in the home of the First Nations sample consisted of English, Ojibway, Cree, Ojicree, and French. Although the majority of the First Nations sample spoke English, English proficiency may play an important role in accounting for differences in item responses. The First Nations parents may have misinterpreted the item questions and consequently misrepresented the child’s emotional and behavioural symptoms. Similarly, variations in item responses between the ethnic groups may be attributed to how the parents are interpreting the child’s mental health symptoms. It is plausible that the cultural differences between parents may influence how the parents endorse similar symptoms. The variation in item responses may be credited to different interpretations of the child’s emotional and behavioural symptoms. Consequently, resulting in different response styles that may influence the endorsement of items as the severity of the mental health

symptoms may vary among the cultures. Additional research is warranted to gain a clearer understanding of item response differences between First Nations and non-First Nations children and adolescents.

It is also important to note that cultures are not homogenous static systems, but rather cultures are “constantly in flux” (Kirmayer, Simpson, & Cargo, 2003). Cultural groups may differ on the degree of affiliation with the dominant cultural context (Dillard & Manson, 2000). Within a particular cultural group there are significant discrepancies in attitudes, practice, and knowledge among the individuals, in addition to varying degrees of contestation, conflict, and resistance with the dominant culture beliefs (Kirmayer et al., 2003). Among First Nations people, there are individuals who subscribe to First Nations cultures as the culture presently exists and there are also individuals who subscribe to the dominant Western culture to various degrees (Waldram, 1997). Williams and Ellison (1996) describe the American Indian culture on a continuum, with four styles of living: traditional, marginal, middle-class, and pan-Indian.

First Nation individuals who subscribe to traditional living adhere to the cultural values and customs. The values may include harmony among individuals and the environment and may encompass various ceremonies and rituals for healing. First Nations individuals who follow marginal living subscribe to traditional culture and dominant culture values. First Nations individuals who subscribe to middle-class values are likely to adhere to the dominant Western values. Lastly, First Nations individuals who subscribe to pan-Indian culture, strive to re-establish lost traditional customs and values (Williams & Ellison, 1996). Consequently, would the cultural subscription of the First Nations parent influence their responses on the BCFPI?

One could speculate that the parents' cultural subscription to either the traditional, marginal, middle-class, or pan-Indian may influence their responses to the mental health questions on the BCFPI. Tseng (2001) acknowledges that culture largely influences an individual's understanding, perception, and interpretation of mental health symptoms. Thus, one may hypothesize that an individual's parent's cultural subscription, may impact their understanding, perception, and interpretation of their children's emotional and behavioural symptoms. These perceptions will ultimately vary based upon whether the parent subscribes to traditional, marginal, middle-class, or pan-Indian lifestyle. Consequently, this may result in different interpretations of the child's emotional and behavioural symptoms and the severity of the child's presentation.

To summarize, the present study examined differential item and test functioning among First Nations and non-First Nations children and adolescents on the mental health scales of the BCFPI. The two ethnic groups demonstrated differences on most items of the mental health scales. However, these differences are eliminated when the items were aggregated into scale scores. The bias is reduced at the scale level due to the cancellation process of the nuisance determinants. Hence, the differential item functioning or bias disappears when considering group differences at the scale level.

Clinical Implications

The clinical implications of these research findings are pivotal. The analyses addressed bias at the item and scale score levels for the mental health scales of the BCFPI and have important implications for clinical practice. The results suggest item responses differences between First Nations and non-First Nations children and

adolescents on all mental health scales. Consequently, due to differential item functioning or bias at the item level, item responses on the mental health scales should not be interpreted for First Nations children and adolescents. However, the bias at the item level was not sustained at the scale level, suggesting that the scale scores of the BCFPI can be utilized among First Nations children and adolescents within Northwestern Ontario in an unbiased manner. The results of the BCFPI parent reports are valid measures of childhood emotional and behavioural disorders. The mental health scale scores are an accurate interpretation of childhood psychopathology as measured by the BCFPI. This is paramount as the reliability and validity of the BCFPI has not been established among First Nations children and adolescents. Consequently, the scale scores of the BCFPI can be utilized among First Nations children and adolescents without any fear of measurement bias.

Limitations

This research study is not without limitations. The First Nations sample included in the present study was from an agency located in Northwestern Ontario that served primarily Ojibway, Cree, and Oji-Cree individuals. As mentioned previously, there is a large cultural diversity among Canada's First Nations people. It would be erroneous to conclude that the research findings from this study can be generalized to all of Canada's First Nations people. The various bands have their own language, beliefs, values, and traditions. Therefore, these research findings are limited to First Nations children and adolescents located within Northwestern Ontario. It would be advantageous for future research to examine differential item functioning and differential test functioning among First Nations children and adolescents in other geographical areas within Ontario.

Another limitation of the analyses is that 10% of the First Nations participants were in foster care. It is unknown how many of these children were placed with First Nations foster parents. It is plausible that the First Nations children may have been placed in a non-First Nations foster home and the foster parent would have responded to the BCFPI assessment. The cultural identity of the BCFPI respondents is an important factor, as discussed above. Moreover, it is also important to acknowledge that the 10% of children in foster care may not be an accurate representation. The actual number of children in foster care may be higher, as Davidson-Arad (2005) notes that outcome studies have indicated high rates of psychological, behavioral, social, and education problems among in children in foster care. Consequently, one would expect the number of children in foster care to be higher. Perhaps the low percentage of children in foster care may be attributed to incorrect information regarding whether the parent or foster parent completed the BCFPI.

A further limitation of the present study is that the study only included the parent reports of the BCFPI; the youth and teachers versions of the BCFPI were not included in the analyses. Cunningham et al. (2001) speculates that the parent, teacher, and youth reports would differ as each report has a different perspective on the child's emotional and behavioural symptoms. Consequently, despite the lack of differential test functioning on the parent reports of the BCFPI among First Nations and non-First Nations children and adolescents, it is conceivable that bias may exist among the scale level of the teacher and youth reports of the BCFPI. Hence, future research should examine item bias and test bias between First Nations and non-First Nations children and adolescents on the teacher and youth reports of the BCFPI.

Lastly, Cunningham et al. (2001) indicates that despite the BCFPI is based upon DSM-IV diagnostic criteria, the BCFPI is not diagnostic, but rather a screening tool. The authors indicate that additional information is required for a diagnosis of emotional or behavioural disorders, as the BCFPI is not a comprehensive assessment. The BCFPI fails to assess for all childhood disorders outlined in the DSM-IV, such as tourette's syndrome, obsessive-compulsive disorder, or thought disorders (Cunningham et al., 2001). As a result, the results from the present study cannot be generalized to other diagnostic categories not assessed by the BCFPI.

Conclusion

The present research study utilized differential item functioning and differential test functioning to assess bias on the mental health scales of the BCFPI between First Nations and non-First Nations children and adolescents. The results of the study support the use of the BCFPI among First Nations children and adolescents. Despite differences between First Nations and non-First Nations children and adolescents at the item level, these differences are not sustained at the scale level. Consequently, the scale scores are comparable between the two ethnic groups and the BCFPI is an accurate assessment of emotional and behaviour symptoms among First Nations children and adolescents. This analysis is paramount as mental health agencies that serve children and adolescents across the province of Ontario are government mandated to implement the use of the BCFPI. Although these research findings are critical, additional research is warranted, as there is a lack of research that has examined mental health issues in First Nations children and adolescents.

References

- Allen, J. (1998). Personality assessment with American Indians and Alaska Natives: Instrument considerations and service delivery style. *Journal of Personality Assessment*, 70, 17-42.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders. Fourth edition, text revision*. Washington, DC: American Psychiatric Press.
- Assembly of First Nations. (1994). *Breaking the silence: An interpretive study of residential school impact and healing as illustrated by the stories of First Nations Individuals*.
- Baker, F. B. (1993). *Equate computer program version 2.1*.
- Barwick, M. A., Schmidt, J., & Hodges, K. (2004). *Culturally competent evaluation: Clinical considerations for rating the Child Adolescent Functional Assessment Scale (CAFAS) with Aboriginal children and youth*. Toronto, Ontario: The Hospital for Sick Children.
- Beals, J., Piasecki, J., Nelson, S., Jones, M., Keane, E., Dauphinais, P., Red Shirt, R., Sack, W., H., & Manson, S. M. (1997). Psychiatric disorder among American Indian Adolescents: Prevalence in Northern Plains Youth. *Journal of American Academy of Child and Adolescent Psychiatry*, 36, 1252-1259.
- Bejar, I. I. (1983). Introduction to item response models and their assumptions. In R. K. Hambleton (Ed.), *Applications of item response theory* (pp. 1-23). Vancouver, BC: Educational Research Institute of British Columbia.
- Bonder, B., Martin, L., & Miracle, A. (2002). *Culture in clinical care*.

Thorofare, NJ: Slack Inc.

- Borzecki, M., Wormith, J. S., & Black, W. H. (1988). An examinations of differences between Native and non-Native psychiatric offenders on the MMPI. *Canadian Journal of Behavioural Science*, 2, 287-301.
- Boyle, M. H., Offord, D. R., Hoffmann, H. G., Catlin, G. P., Byles, J. A., Cadman, D. T., Crawford, J. W., Links, P. S., Rae-Grant, N. I., & Szatmari, P. (1987). Ontario child health study: I. Methodology. *Archives of General Psychiatry*, 44, 826-831.
- Boyle, M. H., Offord, D. R., Racine, Y., Fleming, J. E., Szatmari, P., & Sanford, M. (1993a). Evaluation of the revised Ontario child health study scales. *Journal of Child Psychology and Psychiatry*, 34, 189-213.
- Boyle, M. H., Offord, D. R., Racine, Y., Szatmari, P., & Fleming, J. E. (1993b). Evaluation of the original child health study scales. *Canadian Journal of Psychiatry*, 38, 397-405.
- Butcher, J. N., Braswell, L., & Raney, D. (1983). A cross-cultural comparison of American Indian, Black, and White inpatients on the MMPI and presenting symptoms. *Journal of Consulting and Clinical Psychology*, 51, 587-594.
- Butcher, J. N., Nezami, E., & Exner, J. (1998). Psychological assessment of people in diverse cultures. In S. S. Kazarian and D. R. Evans (Eds.), *Cultural clinical psychology: Theory, research, and practice*. New York, NY: Oxford University Press.
- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: Sage.

- Cuellar, I. (2000). Acculturation and mental health: Ecological transactional relations of adjustment. In I. Cuellar and F. A. Paniagua (Eds.), *Handbook of multicultural mental health*. New York, NY: Academic Press.
- Cunningham, C. E., Pettingill, P., & Boyle, M. (2001). *The Brief Child and Family Phone Interview: Interviewers Manual*. Hamilton Health Sciences, ON: Canadian centre for the study of children at risk.
- Dalrymple, A., O'Doherty, J. J., & Nietschei, K. M. (1995). Comparative analysis of Native admissions and registrations to north-western Ontario treatment facilities: Hospital and community sectors. *Canadian Journal of Psychiatry*, 40, 467-473.
- Davidson-Arad, B. (2005). Fifteen-month follow-up of children at risk: Comparison of the quality of life of children removed from home and children remaining at home. *Children and Youth Services Review*, 27, 1 – 20.
- Dillard, D. A., & Manson, S. M. (2000). Assessing and treating American Indians and Alaska Natives. In I. Cuellar and F. A. Paniagua (Eds.), *Handbook of multicultural mental health*. New York, NY: Academic Press.
- du Toit, M. (2003). *IRT from SSI: Bilog-MG, Multilog, Parascale, Testfact*. Lincolnwood, IL: Scientific Software International Inc.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologist*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Flowers, C. P., Oshima, T. C., & Raju, N. (1999). A description and demonstration of the polytomous-DFIT framework. *Applied Psychological Measurement*, 23, 309–26.

- Frideres, J. S. (1993). *Native peoples in Canada: Contemporary conflicts*. Scarborough, Ontario: Prentice-Hall.
- Gfeller, B. M., & Hundleby, J. D. (1995). Patterns of drug use among Native and white adolescents. *Canadian Journal of Public Health*, 86, 95-97.
- Gotowiec, A., & Beiser, M. (1993). Aboriginal children's mental health: Unique challenges. *Canada's Mental Health*, 7-11.
- Graham, J. R. (1993). *MMPI-2: Assessing personality and psychopathology*. New York: Oxford University Press.
- Haig-Brown, C. (1988). *Resistance & renewal: Surviving the Indian residential school*. Vancouver: Tillacum Library.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Norwell, MA: Kluwer Academic Publishers.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, California: Sage Publications, Inc.
- Herreid, C. F., & Herreid, J. R. (1966). Differences in MMPI scores in Native and non-Native Alaskans. *Journal of Social Psychology*, 70, 191-198.
- Hoffman, T., Dana, R. H., & Bolton, B. (1985). Measured acculturation and MMPI-168 performance of Native American adults. *Journal of Cross-Cultural Psychology*, 16, 243-256.
- Holland, P. W., Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 129-145). Hillsdale, NJ: Erlbaum.
- Hunsley, J. (2004). *Workshop on developing an evidence-based psychological practice*. Thunder Bay, Ontario.

- Hunsley, J., Lee, C. M., & Wood, J. M. (2003). Controversial and questionable assessment techniques. In S. O. Lilienfeld, J. M. Lohr, & S. J. Lynn (Eds.), *Science and pseudoscience in contemporary clinical psychology*. New York: Guilford.
- Ironson, G. H. (1983). Using item response theory to measure bias. In R. K. Hambleton (Ed.), *Applications of item response theory* (155-174). Vancouver, BC: Educational Research Institute of British Columbia.
- Kim, Y., Pilkonis, P. A., Frank, E., Thase, M. E., & Reynolds, C. F. (2002). Differential functioning of the Beck Depression Inventory in late-life patients: Use of item response theory. *Psychology and Aging, 17*, 379-391.
- Kinzie, J. D., Leung, P. K., Boehnlein, J., Matsunaga, D., Johnson, R., Manson, S. M., Shore, J. H., Heinz, J., Williams, M. (1992). Psychiatric epidemiology of an Indian village: A 19-year replication study. *Journal of Nervous and Mental Disease, 180*, 33-39.
- Kirmayer, L. J., Brass, G. M., & Tait, C. L. (2000). The mental health of Aboriginal peoples: Transformations of identity and community. *Canadian Journal of Psychiatry, 45*, 607-616.
- Kirmayer, L., Simpson, C., & Cargo, M. (2003). Healing traditions: Culture, community, and mental health promotion with Canadian Aboriginal peoples. *Australasian Psychiatry, 11*, S15 – S23.
- Kline, J. A., Rozyrko, V. V., Flint, G., & Roberts, A. C. (1973). Personality characteristics of male Native American alcoholic patients. *The International Journal of the Addictions, 8*, 729-732.
- Jenkins, J. H., & Karno, M. (1992). The meaning of expressed emotion:

- Theoretical issues raised by cross-cultural research. *American Journal of Psychiatry*, 149, 9-21.
- Lalinec-Michaud, M., Subak, M. E., Ghadirian, A. M., & Kovess, V. (1991). Substance misuse among Native and rural high school students in Quebec. *Internal Journal of Addiction*, 26, 1003-1012.
- Lilienfeld, S., Lynn, S., & Lohr, J. (Eds.) (2003). *Science and pseudoscience in clinical psychology*. New York: Guilford Press.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Menlo Park, California: Addison-Wesley Publishing Company.
- MacMillan, H., Walsh, C., Jamieson, E., Crawford, A., & Boyle, M. (2000). *Children's health*. First Nations and Inuit Regional Health Surveys.
- Mandelzys, N., & Lane, E. B. (1980). The validity of the MMPI as it pertains to Canadian Native inmates. *Canadian Journal of Criminology*, 22, 188-196.
- Marsella, A. J., & Yamada, A. M. (2000). Culture and mental health: An introduction and overview of foundations, concepts, and issues. In I. Cuellar and F. A. Paniagua (Eds.), *Handbook of multicultural mental health*. New York, NY: Academic Press.
- Miller, J. R. (1996). *Shingwauk's vision: A history of Native residential schools*. Toronto: University of Toronto Press Inc.
- Millsap, R. E., & Everson, H. T. (1999). Methodology review: Statistical

- approaches for assessing measurement bias. *Applied Psychological Measurement*, 17, 297-334.
- Myers, T., Bullock, S. L., Calzavara, L. M., Cockerill, R., & Marshall, V. W. (1997). Differences in sexual risk-taking behaviour with state of inebriation in an Aboriginal population in Ontario, Canada. *Journal Studies of Alcohol*, 58, 312-322.
- Page, R. D., & Bozlee, S. (1982). A cross-cultural MMPI comparison of alcoholics. *Psychological Reports*, 50, 639-646.
- Pollack, D., & Shore, J. H. (1980). Validity of the MMPI with Native Americans. *American Journal of Psychiatry*, 137, 946-950.
- Raju, N. S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement*, 14, 197-207.
- Raju, N. S., van der Linden, W. J., & Fleer, P. F. (1995). IRT-based internal measures of differential functioning of items and tests. *Applied Psychological Measurement*, 19, 353-368.
- Rosnow, R. L., & Rosenthal, R. (1988). Focused tests of significance and effect size estimation in counselling psychology. *Journal of Counselling Psychology*, 35, 203-208.
- Royal Commission on Aboriginal People. (1992). *Aboriginal peoples in urban centres: Report of the National round table on Aboriginal urban issues*.
- Royal Commission on First Nations People. (1995). *Suicide among Aboriginal People: Royal Commission Report*.
- Roy, C., Choudhuri, A., & Irvine, D. (1970). The prevalence of mental disorders

- among Saskatchewan Indians. *Journal of Cross-Cultural Psychology*, 1, 383-392.
- Sampath, H. M. (1974). Prevalence of psychiatric disorders in a Southern Baffin Island Eskimo settlement. *Canadian Psychiatric Association Journal*, 19, 363-367.
- Santor, D. A., & Coyne, J. C. (2001). Evaluating the continuity of symptomatology between depressed and non-depressed individuals. *Journal of Abnormal Psychology*, 110, 216-225.
- Santor, D. A., & Ramsay, J. O. (1998). Progress in the technology of measurement: Applications of item response models. *Psychological Assessment*, 10, 345-359.
- Santor, D. A., Ramsay, J. O., & Zuroff, D. C. (1994). Nonparametric item analysis of the Beck Depression Inventory: Evaluating gender item bias and response option weights. *Psychological Assessment*, 6, 255-270.
- Shealy, R. T., & Stout, W. F. (1993). An item response theory model for test bias and differential test functioning. In P. W. Holland and H. Wainer (Eds.), *Differential item functioning* (pp. 197-239). Hillsdale, NJ: Erlbaum.
- Shore, J. H., Kinzie, J. D., Hampson, J. L., & Pattison, E. M. (1973). Psychiatric epidemiology of an Indian village. *Psychiatry*, 36, 70-81.
- Shore, J. H., & Manson, S. M. (1981). Cross-cultural studies of depression among American Indians and Alaska Natives. *White Cloud Journal of American Indian/Alaska Native Mental Health*, 2, 5-12.
- Statistics Canada. (2001). *Census of Canada*. Canada: Available Online. <http://www12.statcan.ca/english/census01/home/index.cfm>

- Titley, B. (1986). *A narrow vision: Duncan Campbell Scott and the administration of Indian affairs in Canada*. Vancouver: University of British Columbia Press.
- Tseng, W. S. (2001). *Handbook of cultural psychiatry*. New York, NY: Academic Press.
- Tseng, W. S. (2003). *Clinician's guide to cultural psychiatry*. San Diego, California: Academic Press.
- Uecker, A. E., Boutilier, L. R., & Richardson, E. H. (1980). "Indianism" and MMPI scores of men alcoholics. *Journal of Studies on Alcohol*, 14, 357-362.
- Van de Vijver, F. (2000). The nature of bias. In R. H. Dana (Ed.), *Handbook of cross-cultural and multicultural personality assessment*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Venn, J. (1988). MMPI profiles of Native, Mexican, and Caucasian-American male alcoholics. *Psychological Reports*, 62, 427-432.
- Waldram, J. B. (1997). The Aboriginals peoples of Canada. In I. Al-Issa & M. Tousignant (Eds.), *Ethnicity, immigration, and psychopathology*. New York, New York: Plenum Press.
- Waller, N. G., Thompson, J. S., & Wenk, E. (2000). Using IRT to separate measurement bias from true group differences on homogeneous and heterogeneous scales: An illustration with the MMPI. *Psychological Methods*, 5, 125-146.
- Yates, A. (1987). Current status and future direction of research on the American Indian child. *American Journal of Psychiatry*, 114, 1135-1142.

Table 1.

BCFPI Mental Health Scales and Items.

Emotional and Behavioural Disorders

A. Regulation of Attention, Impulsivity, and Activity

1. Distractible, has trouble sticking to an activity
2. Fails to finish things he starts
3. Has difficulty following directions or instructions
4. Impulsive, acts without stopping to think
5. Jumps from one activity to another
6. Fidgets

B. Cooperativeness

1. Cranky
2. Defiant, talks back to adults
3. Blames others for own mistakes
4. Easily annoyed by others
5. Argues a lot with adults
6. Angry and resentful

C. Conduct

1. Steals things at home
2. Destroys things belonging to others.
3. Engages in vandalism
4. Has broken into a house, building or car
5. Physically attacks people
6. Uses weapons when fighting

D. Separation from Parents

1. Worries bad things will happen to loved ones
2. Worries about being separated from loved ones
3. Scared to sleep without parents nearby
4. Overly upset when leaving loved ones
5. Overly upset while away from loved ones
6. Complains of feeling sick before separating

E. Managing Anxiety

1. Worries about doing better at things
2. Worries about past behaviour
3. Worries about doing the wrong thing
4. Worries about things in the future
5. Is afraid of making mistakes
6. Is overly anxious to please people

Table 1.

BCFPI Mental Health Scales and Items.

F. Managing Mood

1. No interest in usual activities
2. Gets no pleasure from usual activities
3. Has trouble enjoying self
4. Not as happy as other children
5. Feels hopeless
6. Unhappy, sad or depressed
7. Has lost a lot of weight without trying
8. Talks about killing himself/herself
9. Deliberately harms self or attempts suicide

First Nations and BCFPI 90

Table 2

Demographic Information for First Nations and non-First Nations Participants.

Demographics	Native		Non-Native Clinical		Non-Native General Population	
	<i>n</i> = 340		<i>n</i> = 1242		<i>n</i> = 1683	
	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Single Parent						
Yes	151	44.4	823	66.3	1232	73.2
No	164	48.2	351	28.3	357	21.2
Missing	25	7.4	68	5.5	94	5.6
Language Spoken in Home						
English	294	86.5	1157	93.2	1506	89.5
Ojibway	4	1.2	0	0	0	0
Cree	1	0.3	0	0	0	0
Ojicree	6	1.8	0	0	0	0
French	1	0.3	6	0.5	10	0.6
Italian	0	0	4	0.3	3	0.2
Portuguese	0	0	2	0.2	4	0.2
Hindi	0	0	1	0.1	7	0.4
German	0	0	0	0	4	0.2
Ukranian	0	0	0	0	1	0.1
Greek	0	0	0	0	6	0.4
Other	0	0	11	0.9	69	4.1
Missing	34	10	61	4.9	60	3.6

Table 2

Demographic Information for First Nations and non-First Nations Participants.

Level of Education						
No Schooling	4	1.2	0	0	0	0
Some Elementary	17	5	15	1.2	37	2.2
Complete Elementary	25	7.4	34	2.7	69	4.1
Some Secondary	106	31.2	283	22.8	458	27.2
Complete Secondary	34	10	264	21.3	448	26.6
Some College	41	12.1	176	14.2	197	11.7
Complete College	31	9.1	184	14.8	177	10.5
Some University	21	6.2	87	7	85	5.1
Complete University	3	0.9	139	11.2	147	8.7
Missing	58	17.1	60	4.8	65	3.9
Spouse's Level of Education						
No Schooling	7	2.1	0	0	1	0.1
Some Elementary	12	3.5	23	1.9	43	2.6
Complete Elementary	18	5.3	26	2.1	68	4
Some Secondary	72	21.2	204	16.4	352	20.9
Complete Secondary	21	6.2	178	14.3	309	18.4
Some College	18	5.3	89	7.2	119	7.1
Complete College	22	6.5	124	10	139	8.3
Some University	9	2.6	49	3.9	78	4.6
Complete University	4	1.2	131	10.5	134	8
Missing	157	46.2	418	33.7	440	26.1

First Nations and BCFPI 92

Table 2

Demographic Information for First Nations and non-First Nations Participants.

Income Before Taxes						
<\$10,000	54	15.9	61	4.9	92	5.5
\$10,000-14,999	42	12.4	95	7.6	139	8.3
\$15,000-19,999	19	5.6	78	6.3	85	5.1
\$20,000-29,000	42	12.4	147	11.8	161	9.6
\$30,000-39,000	24	7.1	167	13.4	270	16
\$40,000-49,000	11	3.2	207	16.7	325	19.3
\$50,000-59,000	12	3.5	165	13.3	208	12.4
>\$60,000	14	4.1	219	17.6	211	12.5
Missing	122	35.9	103	8.3	192	11.4

Table 3

BCFPI Scale Scores, t Values, and Effect Size for the First Nations Clinical Sample and non-First Nations Clinical Sample.

BCFPI Scale	Total items	Native			Non-Native			t	p	Effect Size
		M	SD	Range	M	SD	Range			
Attention, Impulsivity, & Activity	6	6.41	3.42	0-12	6.41	3.27	0-12	-0.02	0.98	.002
Cooperativeness	6	6.64	3.41	0-12	7.11	3.25	0-12	-2.24	0.025	.099
Conduct	6	2.07	2.12	0-11	1.80	2.10	0-12	2.04	0.04	.088
Separation From Parents	6	3.86	2.78	0-12	3.48	3.05	0-12	2.16	0.03	.090
Managing Anxiety	6	3.57	2.93	0-12	4.49	3.14	0-12	-4.93	0.01	.206
Managing Mood	9	4.01	3.40	0-15	5.81	3.77	0-17	-8.17	0.01	.328

First Nations and BCFPI 94

Table 4

BCFPI Scale Scores, t Values, and Effect Size for the First Nations Clinical Sample and non-First Nations General Population Sample.

BCFPI Scale	Total items	Natives			Non-Natives			t	p	Effect Size
		M	SD	Range	M	SD	Range			
Attention, Impulsivity, & Activity	6	6.41	3.42	0-12	3.20	2.69	0-12	16.16	0.01	.617
Cooperativeness	6	6.64	3.41	0-12	3.24	2.53	0-12	17.25	0.01	.649
Conduct	6	2.07	2.12	0-11	0.35	0.82	0-8	14.54	0.01	.613
Separation From Parents	6	3.86	2.78	0-12	1.88	2.09	0-11	12.24	0.01	.521
Managing Anxiety	6	3.57	2.93	0-12	2.99	2.27	0-12	3.36	0.01	.165
Managing Mood	9	4.01	3.40	0-15	1.59	2.02	0-13	12.30	0.01	.542

Table 5

BCFPI Scale Scores, *t* Values, and Effect Size for the non-First Nations Clinical Sample and non-First Nations General Population Sample.

BCFPI Scale	Total items	Clinical			General Population			<i>t</i>	<i>p</i>	Effect Size
		<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range			
Attention, Impulsivity, & Activity	6	6.41	3.27	0-12	3.20	2.69	0-12	27.63	0.01	.503
Cooperativeness	6	7.11	3.25	0-12	3.24	2.53	0-12	34.24	0.01	.592
Conduct	6	1.80	2.10	0-12	0.35	0.82	0-8	22.34	0.01	.510
Separation From Parents	6	3.48	3.05	0-12	1.88	2.09	0-11	15.52	0.01	.335
Managing Anxiety	6	4.49	3.14	0-12	2.99	2.27	0-12	13.90	0.01	.299
Managing Mood	9	5.81	3.77	0-17	1.59	2.02	0-13	34.68	0.01	.653

Table 6

Differential Item Functioning Among the First Nations Clinical Sample and the non-First Nations Clinical Sample.

<u>BCFPI Scales & Items</u>	<u>Chi Square</u>	<u>P Value</u>
Attention, Impulsivity, Activity		
Distractible, has trouble sticking to an activity	1906.45	.01
Fails to finish things he starts	1364.95	.01
Has difficulty following directions or instructions	1542.37	.01
Impulsive, acts without stopping to think	4366.52	.01
Jumps from one activity to another	64001.89	.01
Fidgets	3779.56	.01
Cooperativeness with Others		
Cranky	40821.30	.01
Defiant, talks back to adults	3286.15	.01
Blames others for own mistakes	11896.96	.01
Easily annoyed by others	2739.05	.01
Argues a lot with adults	5743.63	.01
Angry and resentful	1285.29	.16
Conduct		
Steals things at home	1531.05	.01
Destroys things belonging to others.	2245.05	.01
Engages in vandalism	1528.91	.01
Has broken into a house, building or car	2119.63	.01
Physically attacks people	1236.01	.49
Uses weapons when fighting	1277.09	.20

Table 6

Differential Item Functioning Among the First Nations Clinical Sample and the non-First Nations Clinical Sample.

<u>BCFPI Items and Scales</u>	<u>Chi Square</u>	<u>P Value</u>
Separation From Parents		
Worries bad things will happen to loved ones	20276.67	.01
Worries about being separated from loved ones	1656.62	.01
Scared to sleep without parents nearby	7087.23	.01
Overly upset when leaving loved ones	1985.37	.01
Overly upset while away from loved ones	2872.76	.01
Complains of feeling sick before separating	2259.37	.01
Managing Anxiety		
Worries about doing better at things	15322.80	.01
Worries about past behaviour	6279.12	.01
Worries about doing the wrong thing	13840.86	.01
Worries about things in the future	7647.62	.01
Is afraid of making mistakes	4179.62	.01
Is overly anxious to please people	12715.98	.01
Managing Mood		
No interest in usual activities	1919.38	.01
Gets no pleasure from usual activities	1710.40	.01
Has trouble enjoying self	4663.93	.01
Not as happy as other children	2051.97	.01
Not as happy as other children	1718.84	.01
Unhappy, sad or depressed	6963.65	.01
Has lost a lot of weight without trying	1899.26	.01
Talks about killing himself/herself	1285.67	.17
Deliberately harms self or attempts suicide	1790.18	.01

Table 7

Differential Item Functioning Among the First Nations Clinical Sample and the non-First Nations General Population Sample.

<u>BCFPI Scales & Items</u>	<u>Chi Square</u>	<u>P Value</u>
Attention, Impulsivity, Activity		
Distractible, has trouble sticking to an activity	4854.90	.01
Fails to finish things he starts	4574.63	.01
Has difficulty following directions or instructions	21554.66	.01
Impulsive, acts without stopping to think	18012	.01
Jumps from one activity to another	4995.49	.01
Fidgets	3843.62	.01
Cooperativeness with Others		
Cranky	46046.23	.01
Defiant, talks back to adults	2690.33	.01
Blames others for own mistakes	2018.71	.01
Easily annoyed by others	15765.61	.01
Argues a lot with adults	4539.24	.01
Angry and resentful	8161.52	.01
Conduct		
Steals things at home	5796.59	.01
Destroys things belonging to others.	2220.41	.01
Engages in vandalism	2153.13	.01
Has broken into a house, building or car	4895.49	.01
Physically attacks people	1736.87	.17
Uses weapons when fighting	3819.71	.01

Table 7

Differential Item Functioning Among the First Nations Clinical Sample and the non-First Nations General Population Sample.

<u>BCFPI Items & Scales</u>	<u>Chi Square</u>	<u>P Value</u>
Separation From Parents		
Worries bad things will happen to loved ones	30817.22	.01
Worries about being separated from loved ones	2472.73	.01
Scared to sleep without parents nearby	1840.36	.01
Overly upset when leaving loved ones	1806.06	.02
Overly upset while away from loved ones	3058.12	.01
Complains of feeling sick before separating	1703.45	.34
Managing Anxiety		
Worries about doing better at things	13136.67	.01
Worries about past behaviour	15018.23	.01
Worries about doing the wrong thing	2085.4	.01
Worries about things in the future	3000.76	.01
Is afraid of making mistakes	3401.15	.01
Is overly anxious to please people	109952.4	.01
Managing Mood		
No interest in usual activities	3102.59	.01
Gets no pleasure from usual activities	2302.92	.01
Has trouble enjoying self	1996.39	.01
Not as happy as other children	1768.25	.07
Not as happy as other children	2578.36	.01
Unhappy, sad or depressed	3155.23	.01
Has lost a lot of weight without trying	3729.69	.01
Talks about killing himself/herself	6096.09	.01
Deliberately harms self or attempts suicide	7111.98	.01

First Nations and BCFPI 100

Table 8

Differential Test Functioning Among the First Nations Clinical Sample and the non-First Nations Clinical Sample.

BCFPI Scale	Number of Items	Number of Items with Significant DIF	DTF Cut-off Value	DTF	rDTF
Attention, Impulsivity, & Activity	6	6	.144	.05338	.23104
Cooperativeness	6	5	.144	.00866	.09307
Conduct	6	4	.144	.03145	.17734
Separation from Parents	6	6	.144	.00137	.03701
Managing Anxiety	6	6	.144	.00404	.06353
Managing Mood	9	8	.216	.02088	.14449

Table 9

Differential Test Functioning Among the First Nations Clinical Sample and the non-First Nations General Population Sample.

BCFPI Scale	Number of Items	Number of Items with Significant DIF	DTF Cut-off Value	DTF	rDTF
Attention, Impulsivity, & Activity	6	6	.144	.01608	.12682
Cooperativeness	6	6	.144	.05511	.23475
Conduct	6	5	.144	.00002	.00404
Separation from Parents	6	5	.144	.01561	.12494
Managing Anxiety	6	6	.144	.02463	.15693
Managing Mood	9	8	.216	.00301	.05485

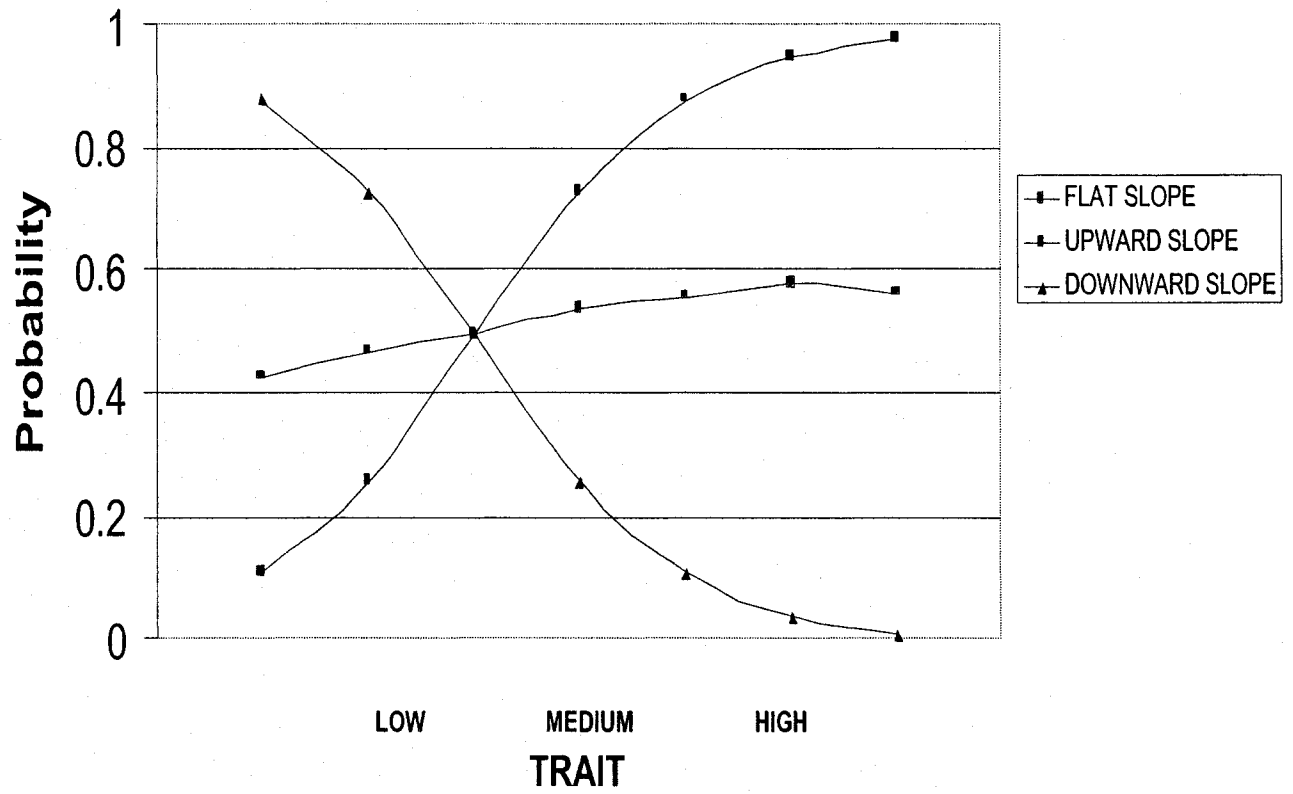


Figure 1.

Flat, Upward, and Downward Slopes.

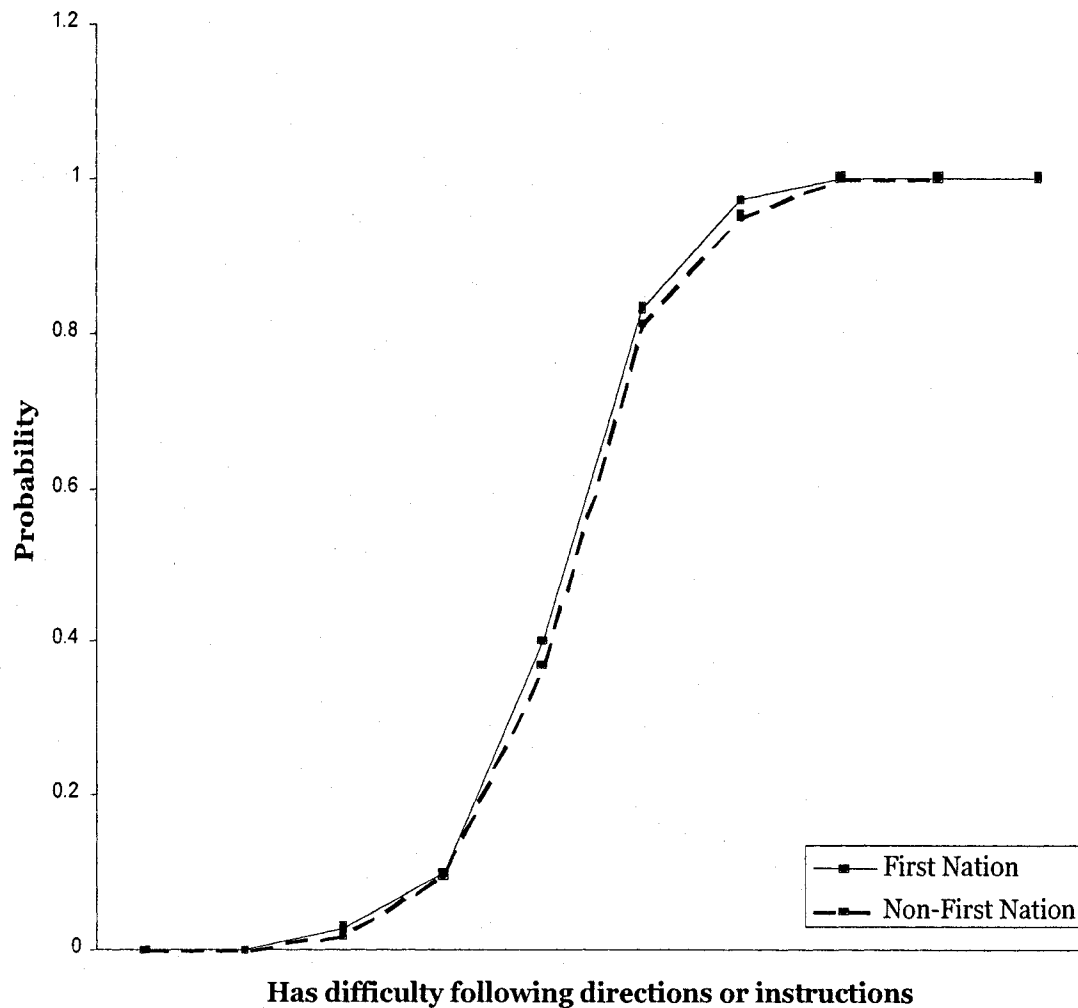


Figure 2.

Example of hypothetical item characteristic curves between First Nations and non-First Nations children and adolescents on the “Has difficulty following directions or instructions” item from the Regulating Attention, Impulsivity, and Activity Level scale of the BCFPI.

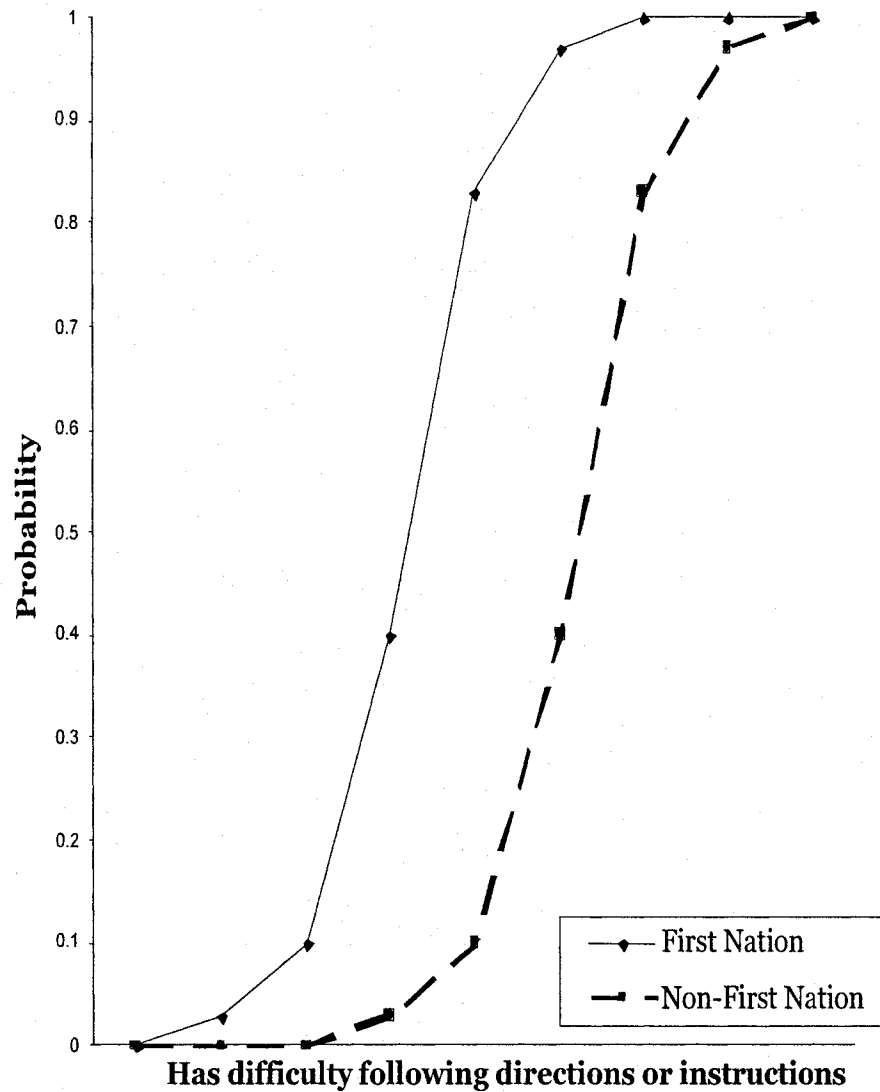


Figure 3.

Example of hypothetical uniform differential item functioning between First Nations and non-First Nations children and adolescents on the "Has difficulty following directions or instructions" item from the Regulating Attention, Impulsivity, and Activity Level scale of the BCFPI.

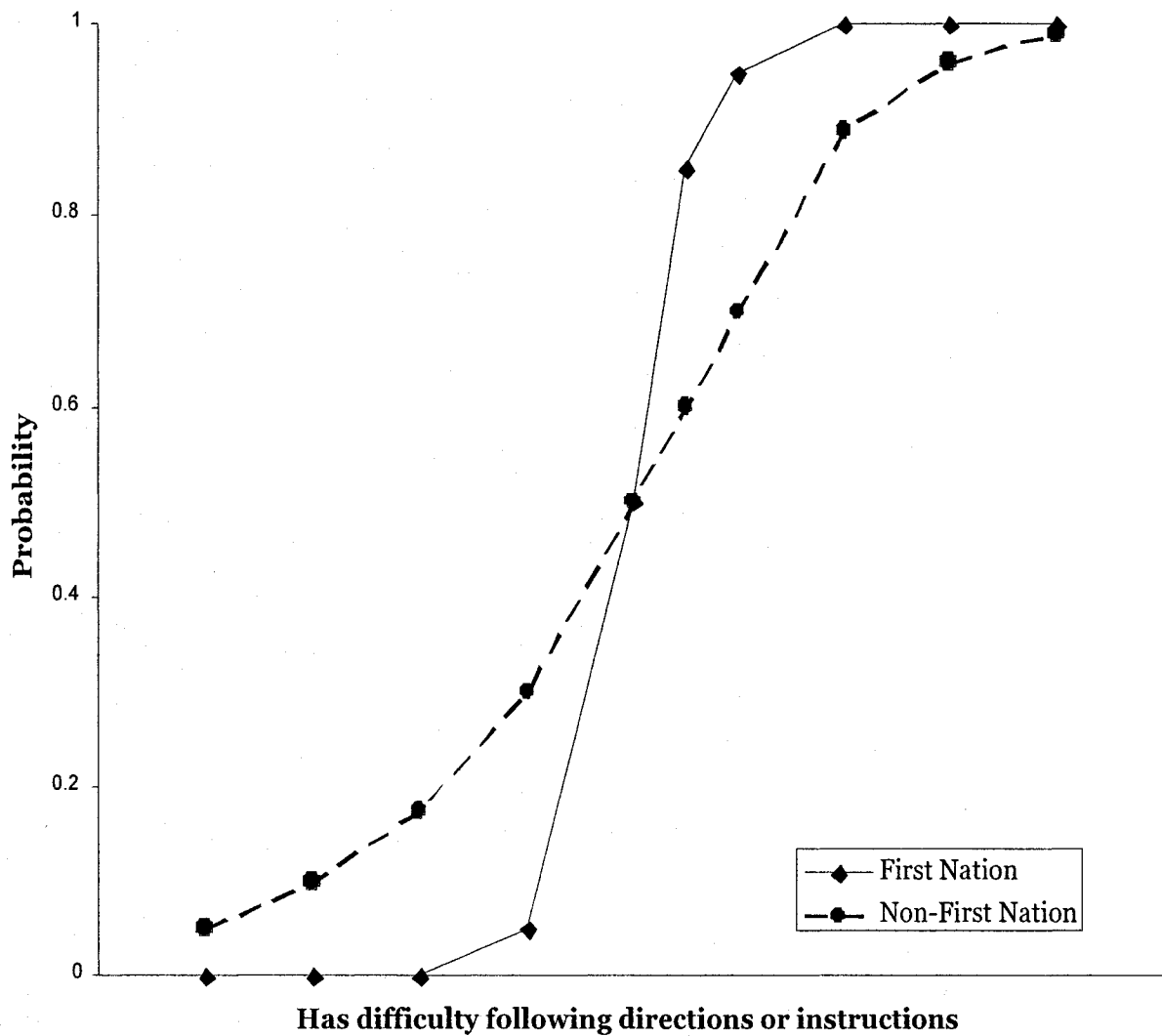


Figure 4.

Example of hypothetical non-uniform differential item functioning between First Nations and non-First Nations children and adolescents on the “Has difficulty following directions or instructions” item from the Regulating Attention, Impulsivity, and Activity Level scale of the BCFPI.

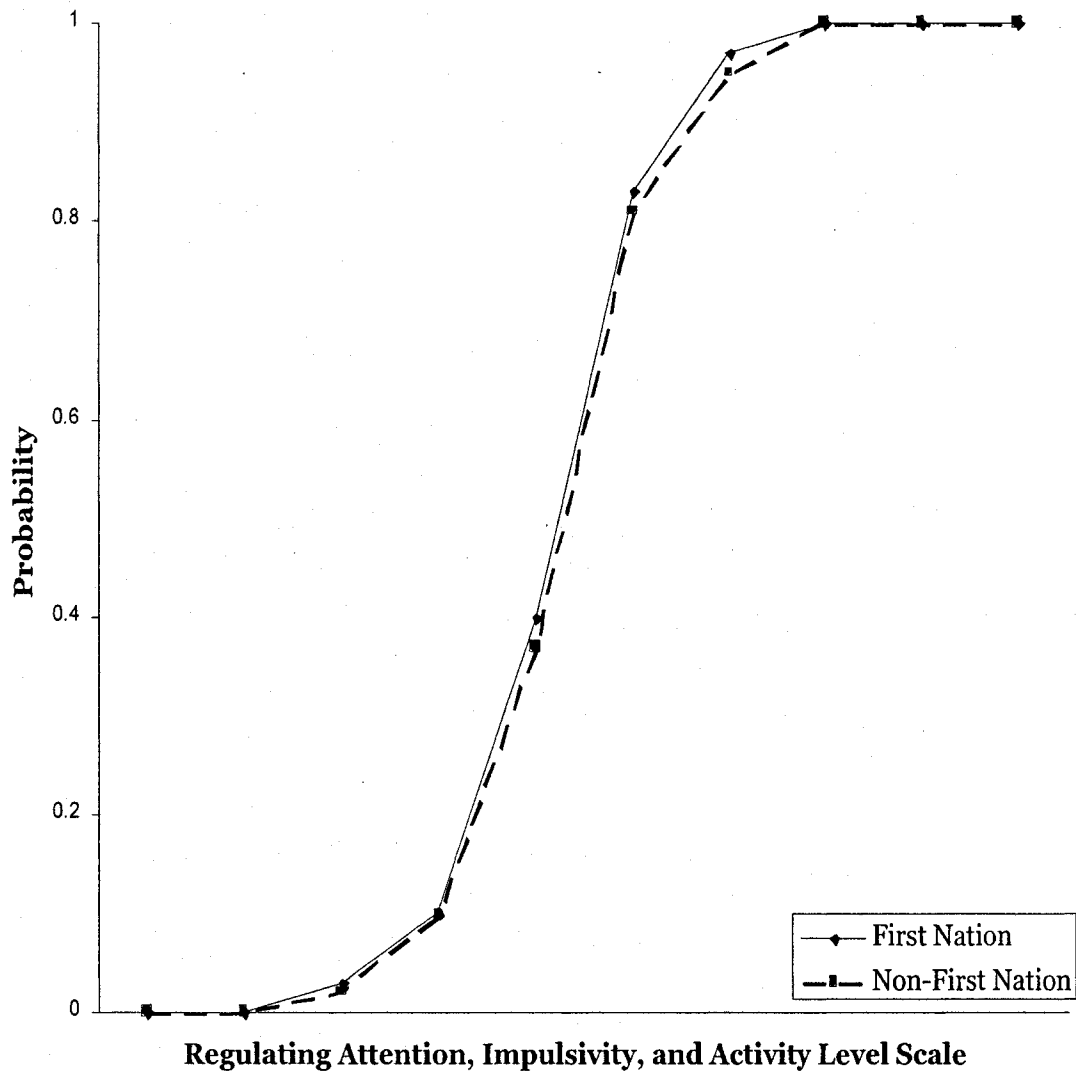


Figure 5.

An illustration of test characteristic curves for no differential test functioning on the Regulation of Attention, Impulsivity, and Activity Level scale from the BCFPI between First Nations and non-First Nations children and adolescents.

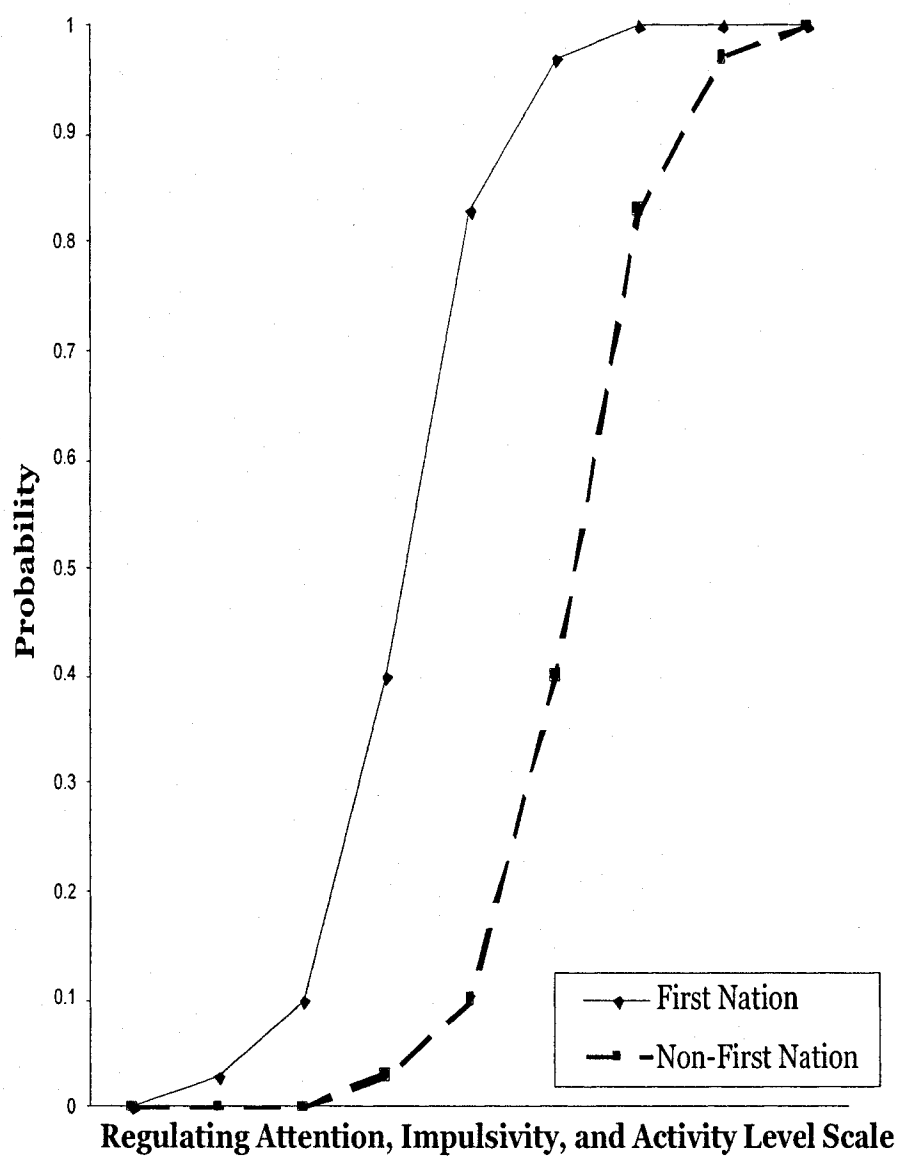


Figure 6.

An illustration of test characteristic curves for differential test functioning on the Regulation of Attention, Impulsivity, and Activity Level scale from the BCFPI between First Nations and non-First Nations children and adolescents

First Nations and BCFPI 108

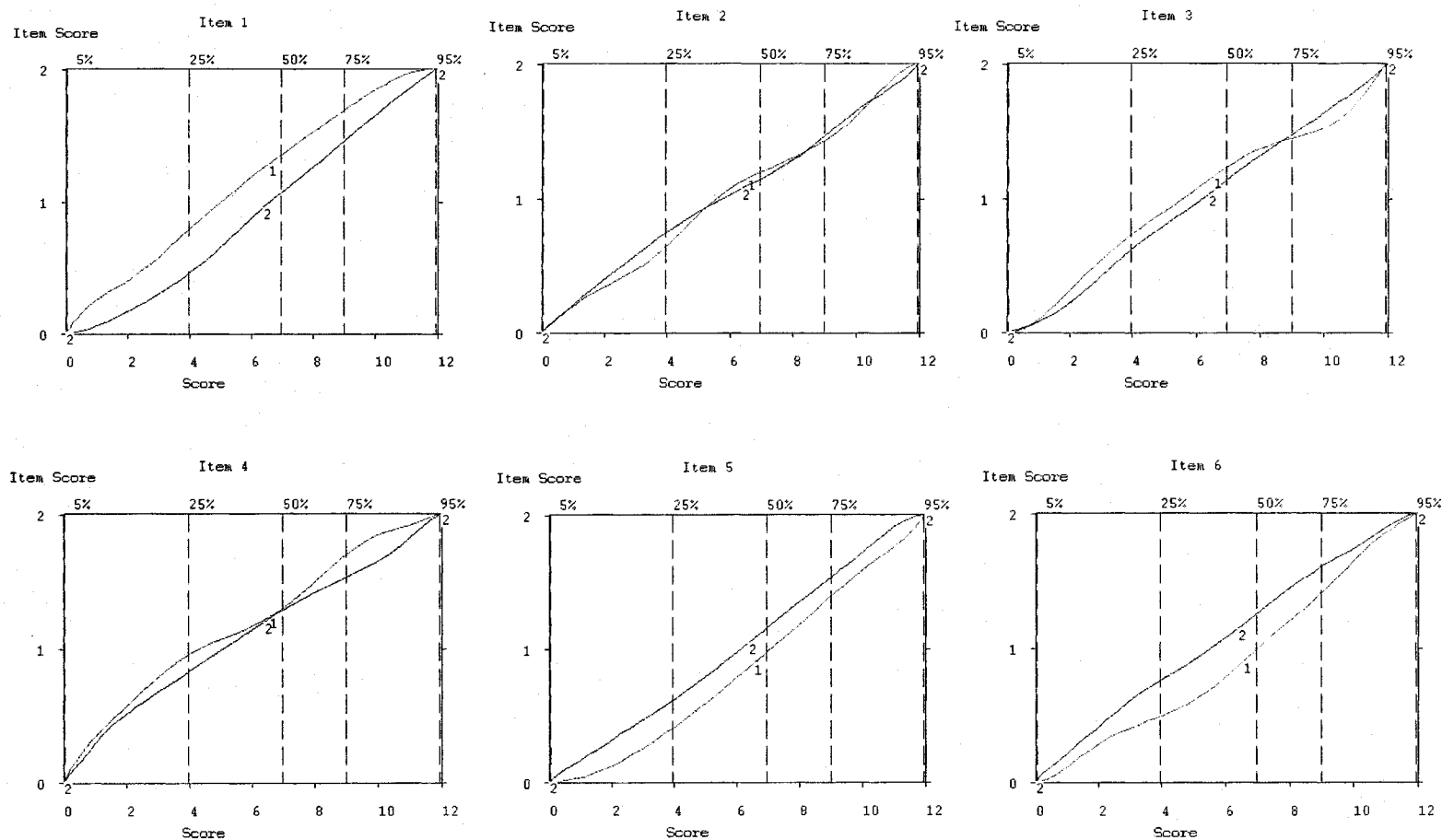


Figure 7.

Item characteristic curves for each item on the Regulation of Attention, Impulsivity, and Activity scale on the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample.

First Nations and BCFPI 109

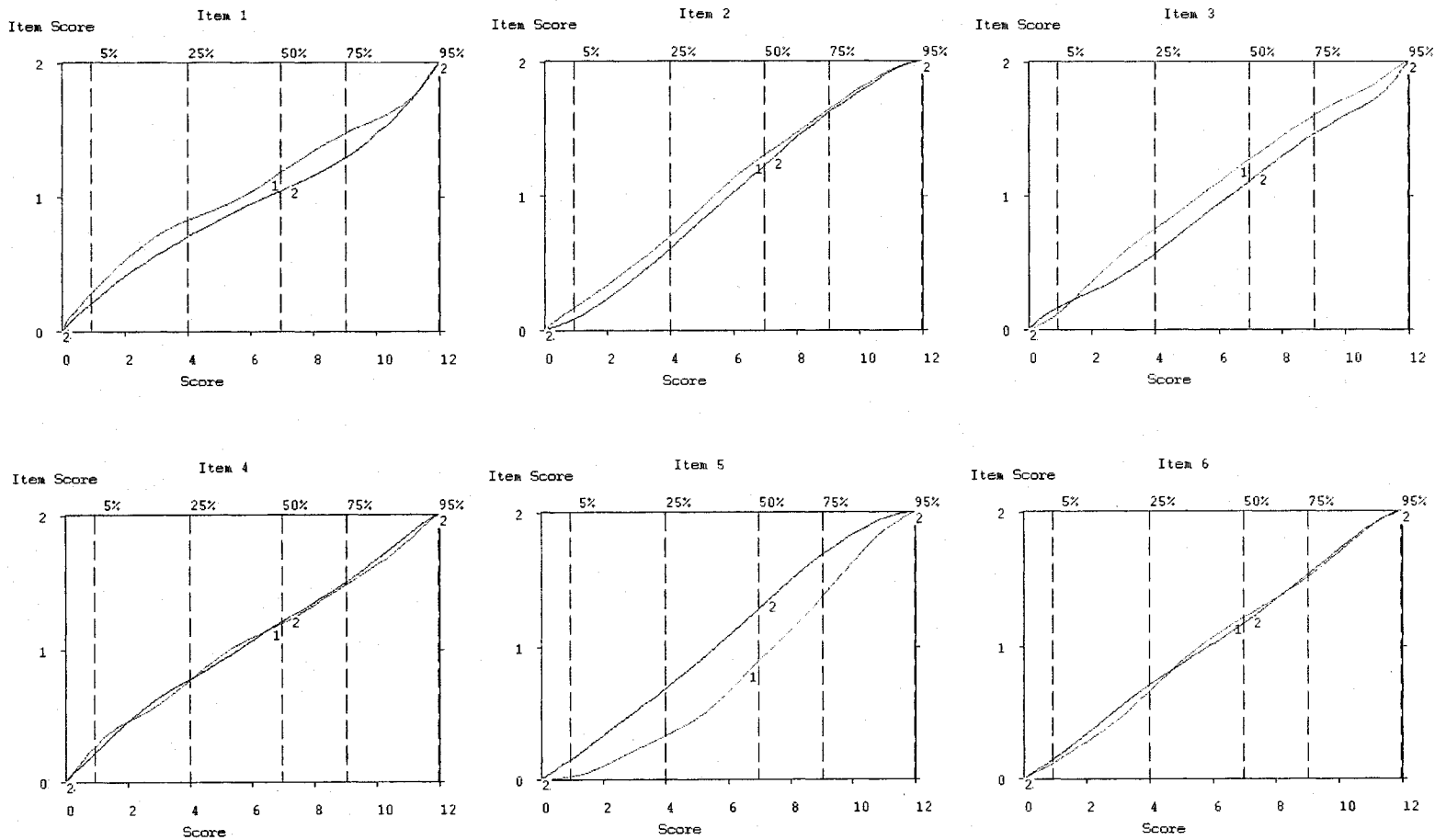


Figure 8.

Item characteristic curves for each item on the Cooperativeness with Others scale on the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample.

First Nations and BCFPI 110

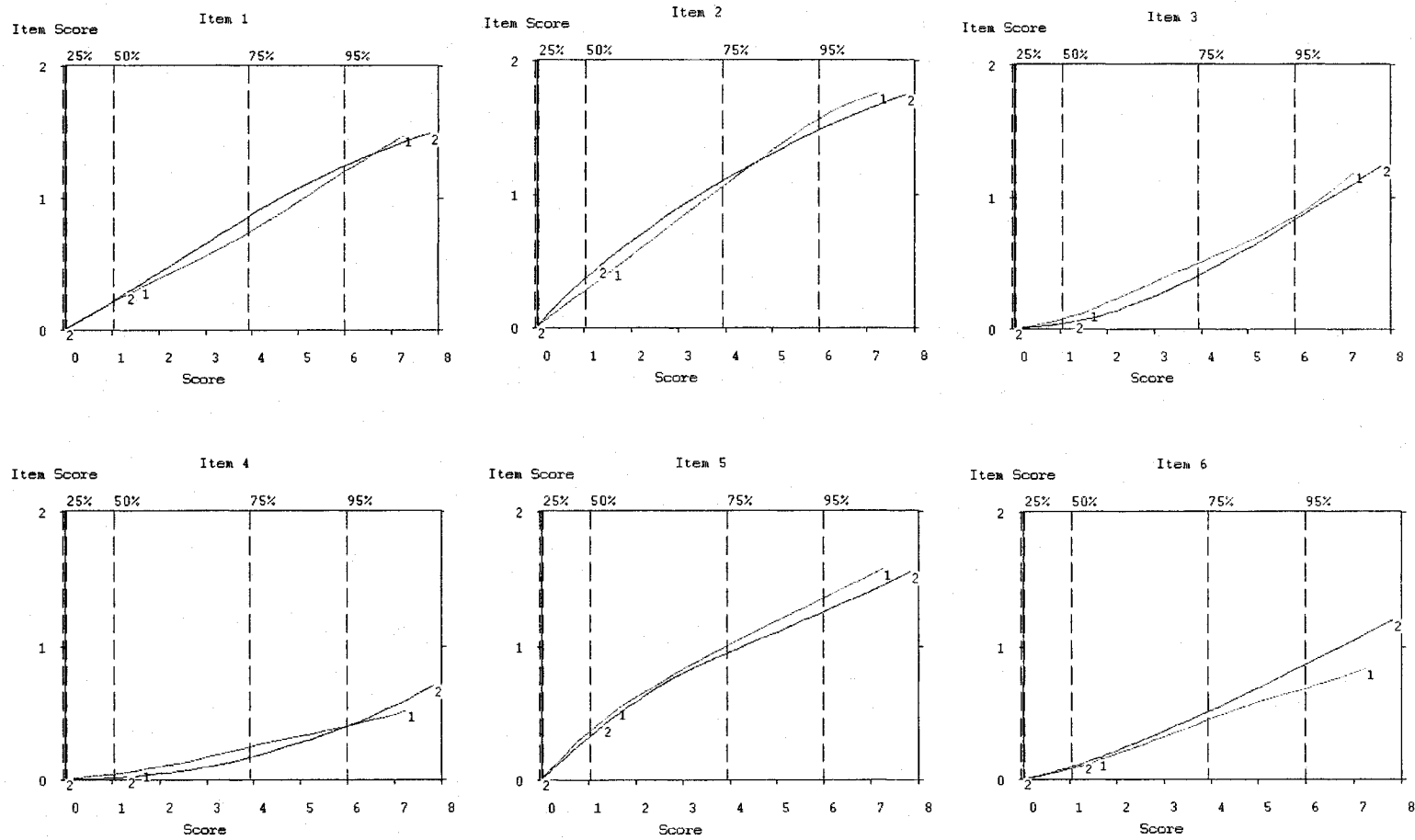


Figure 9.

Item characteristic curves for each item on the Conduct scale on the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample.

First Nations and BCFPI 111

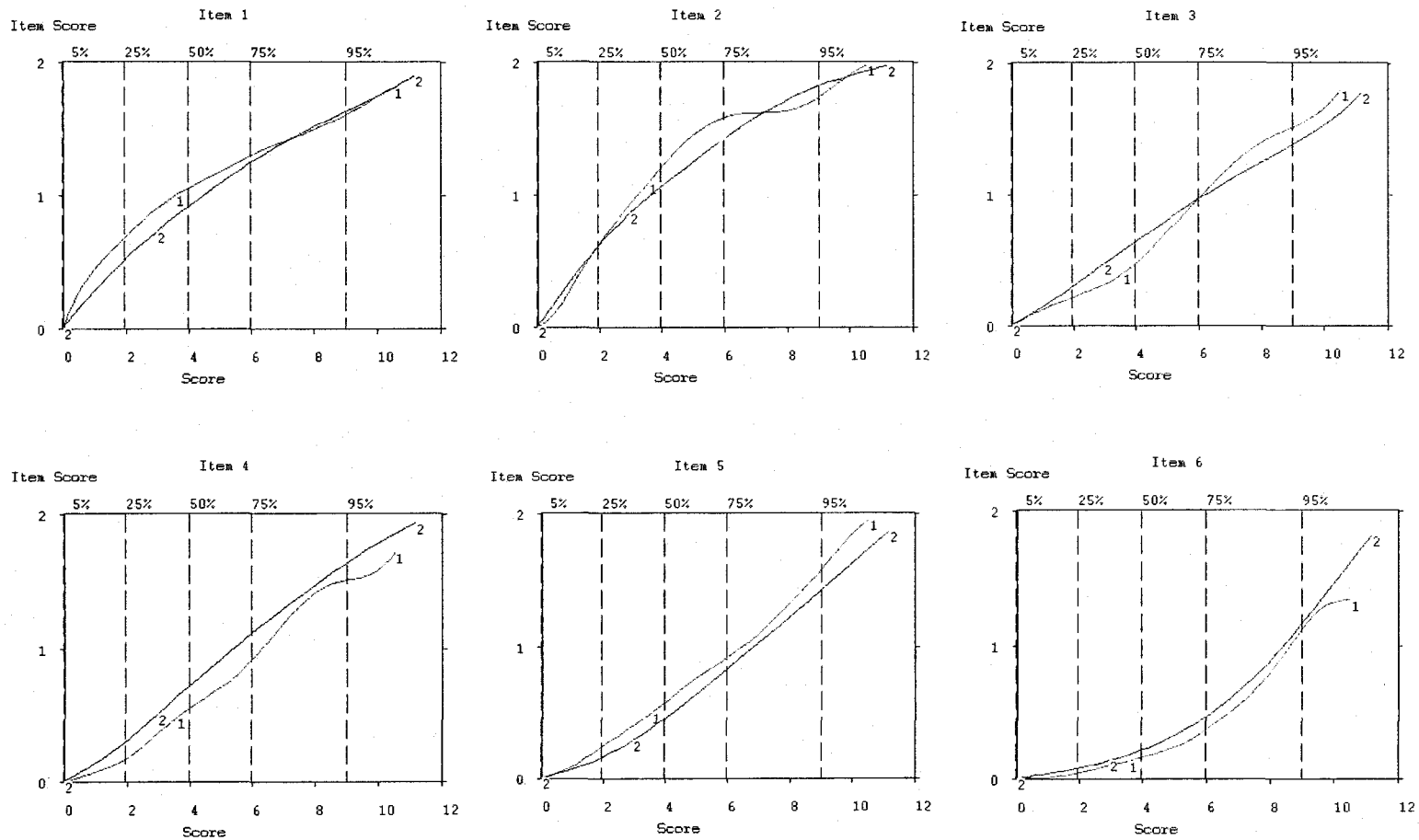


Figure 10.

Item characteristic curves for each item on the Separation from Parents scale on the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample.

First Nations and BCFPI 112

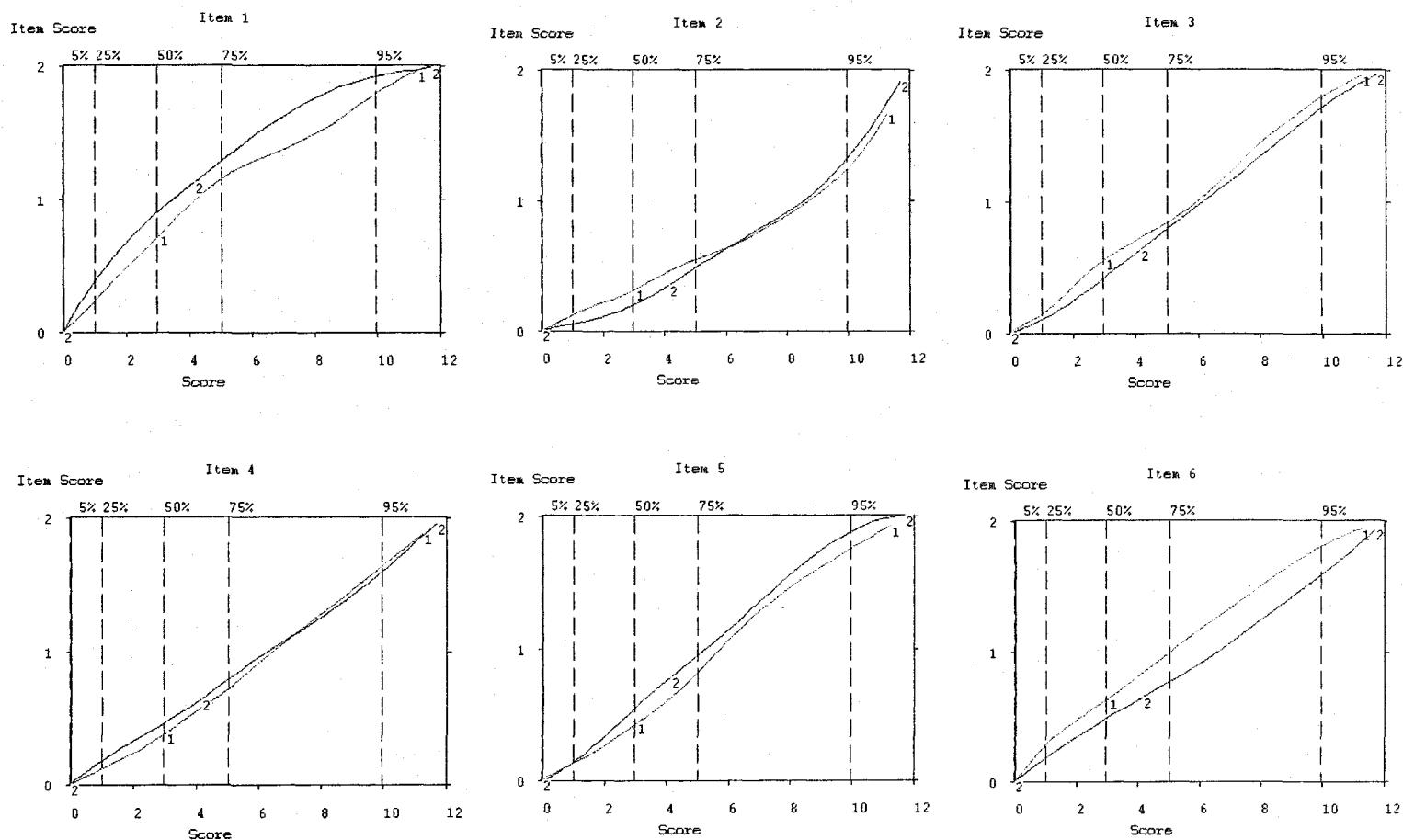


Figure 11.

Item characteristic curves for each item on the Managing Anxiety scale on the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample.

First Nations and BCFPI 113

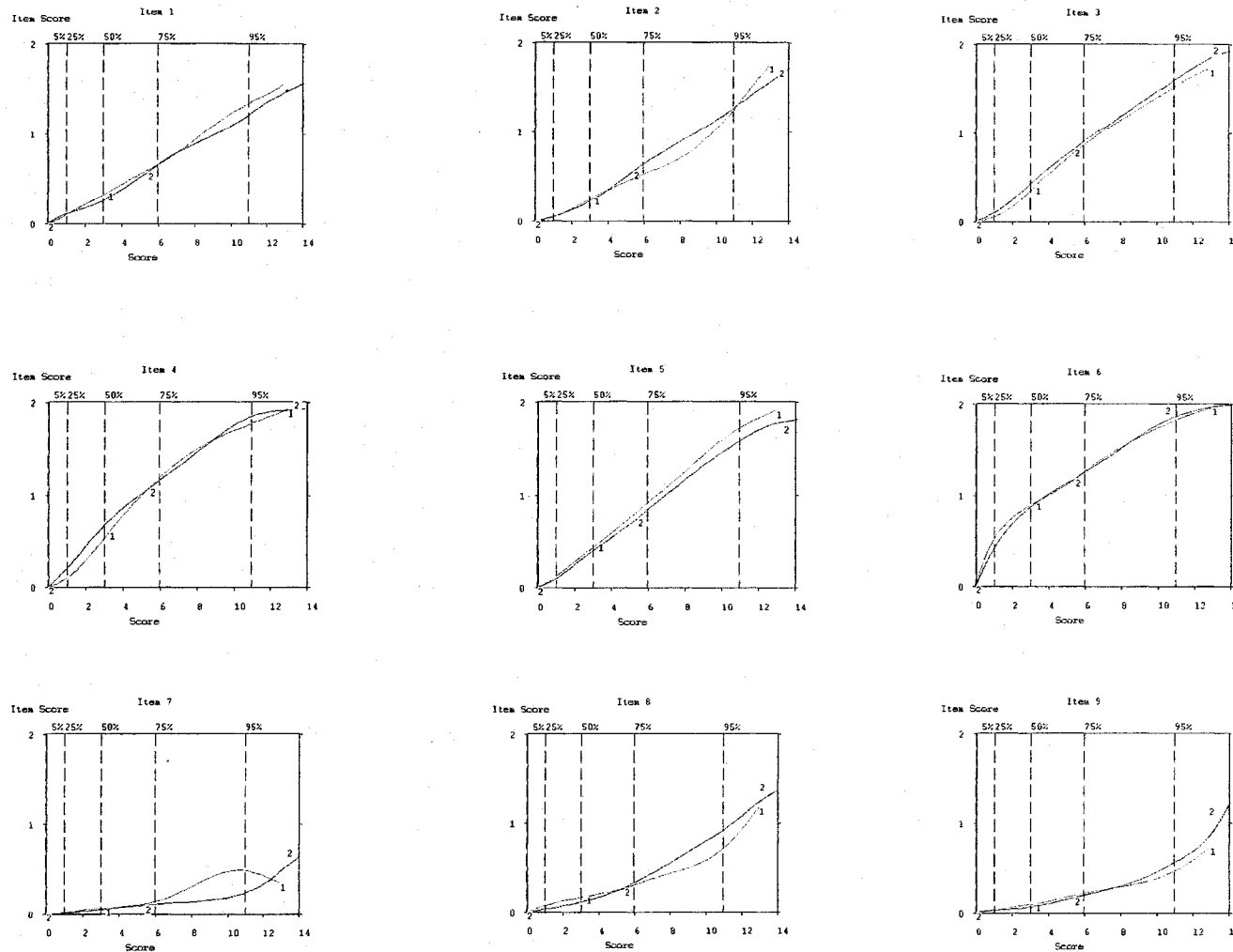


Figure 12.

Item characteristic curves for each item on the Managing Mood scale on the BCFPI among the First Nations clinical sample and the non-First Nations clinical sample.

First Nations and BCFPI 114

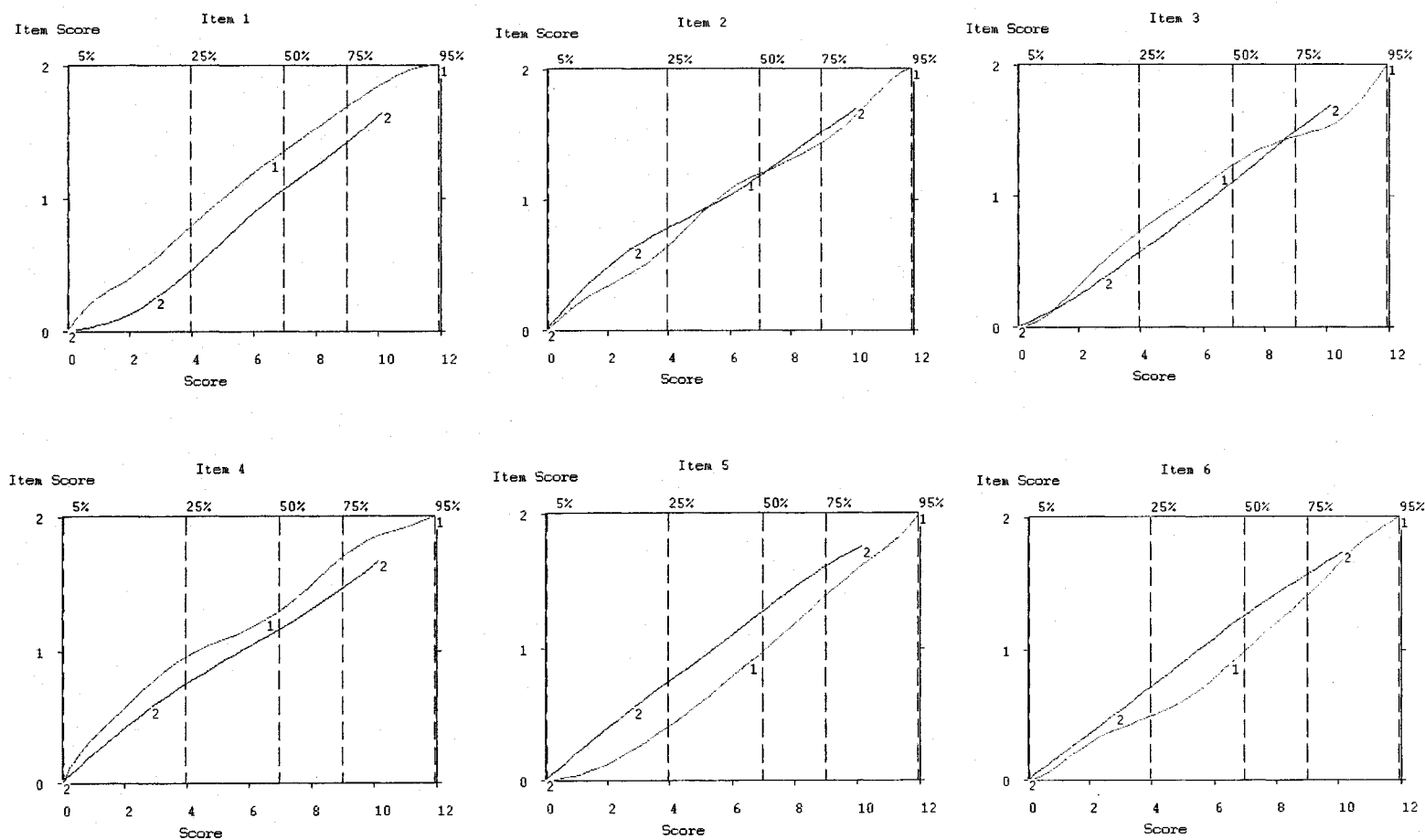


Figure 13.

Item characteristic curves for each item on the Regulating Attention, Impulsivity, and Activity scale on the BCFPI among the First Nations clinical sample and the non-First Nations community sample.

First Nations and BCFPI 115

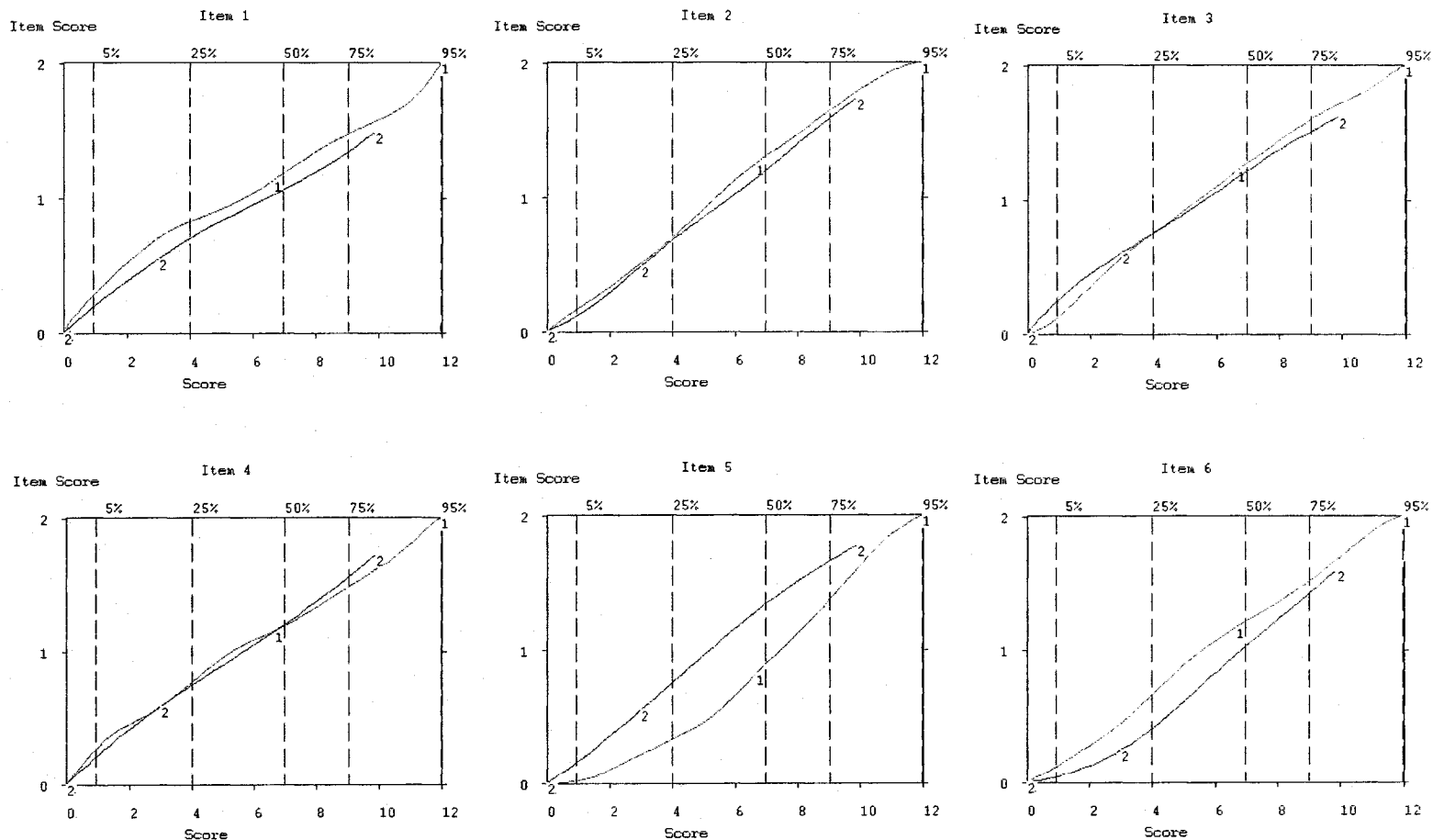


Figure 14.

Item characteristic curves for each item on the Cooperativeness with Others scale on the BCFPI among the First Nations clinical sample and the non-First Nations community sample.

First Nations and BCFPI 116

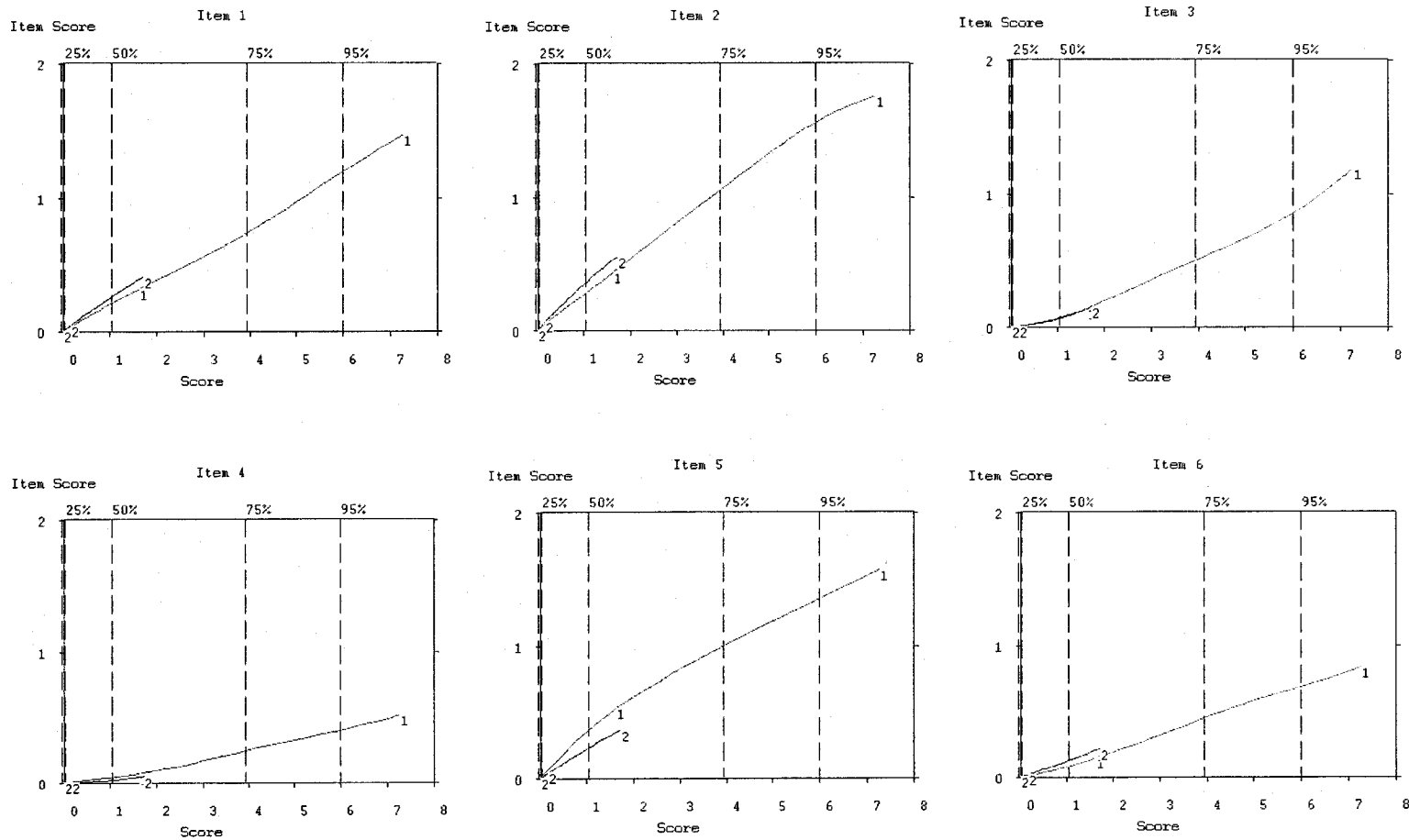


Figure 15.

Item characteristic curves for each item on the Conduct scale on the BCFPI among the First Nations clinical sample and the non-First Nations community sample.

First Nations and BCFPI 117

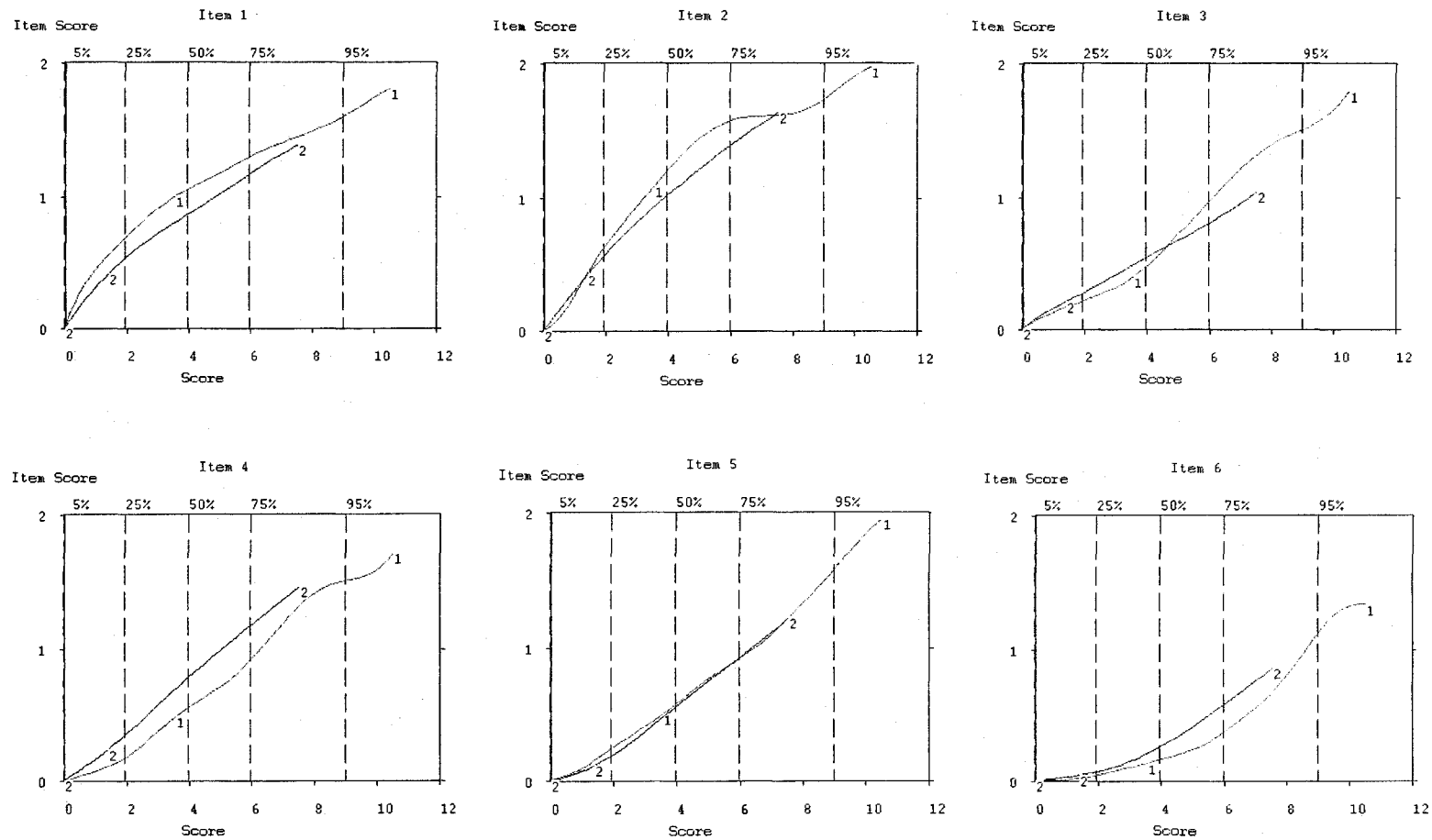


Figure 16.

Item characteristic curves for each item on the Separation from Parents scale on the BCFPI among the First Nations clinical sample and the non-First Nations community sample.

First Nations and BCFPI 118

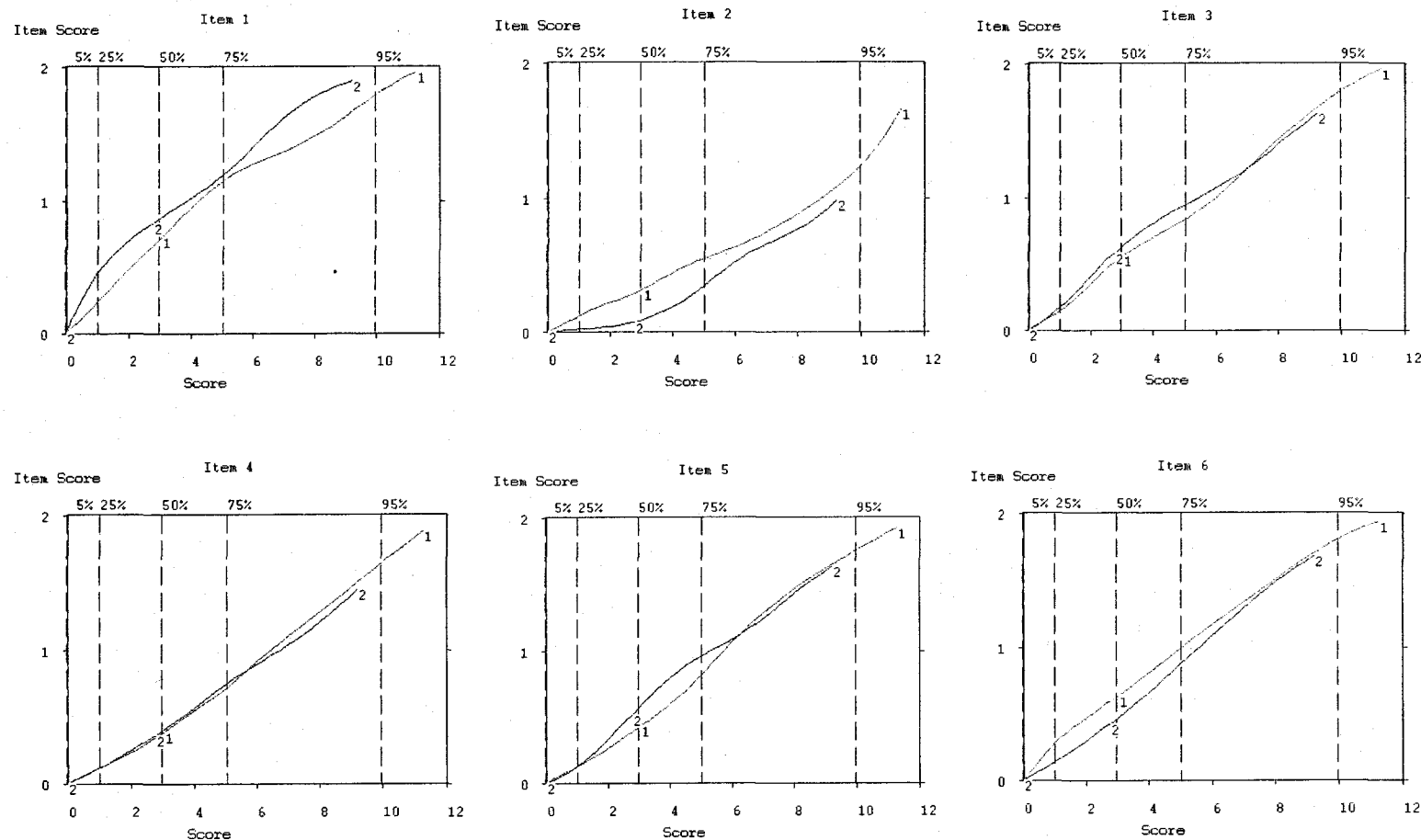


Figure 17.

Item characteristic curves for each item on the Managing Anxiety scale on the BCFPI among the First Nations clinical sample and the non-First Nations community sample.

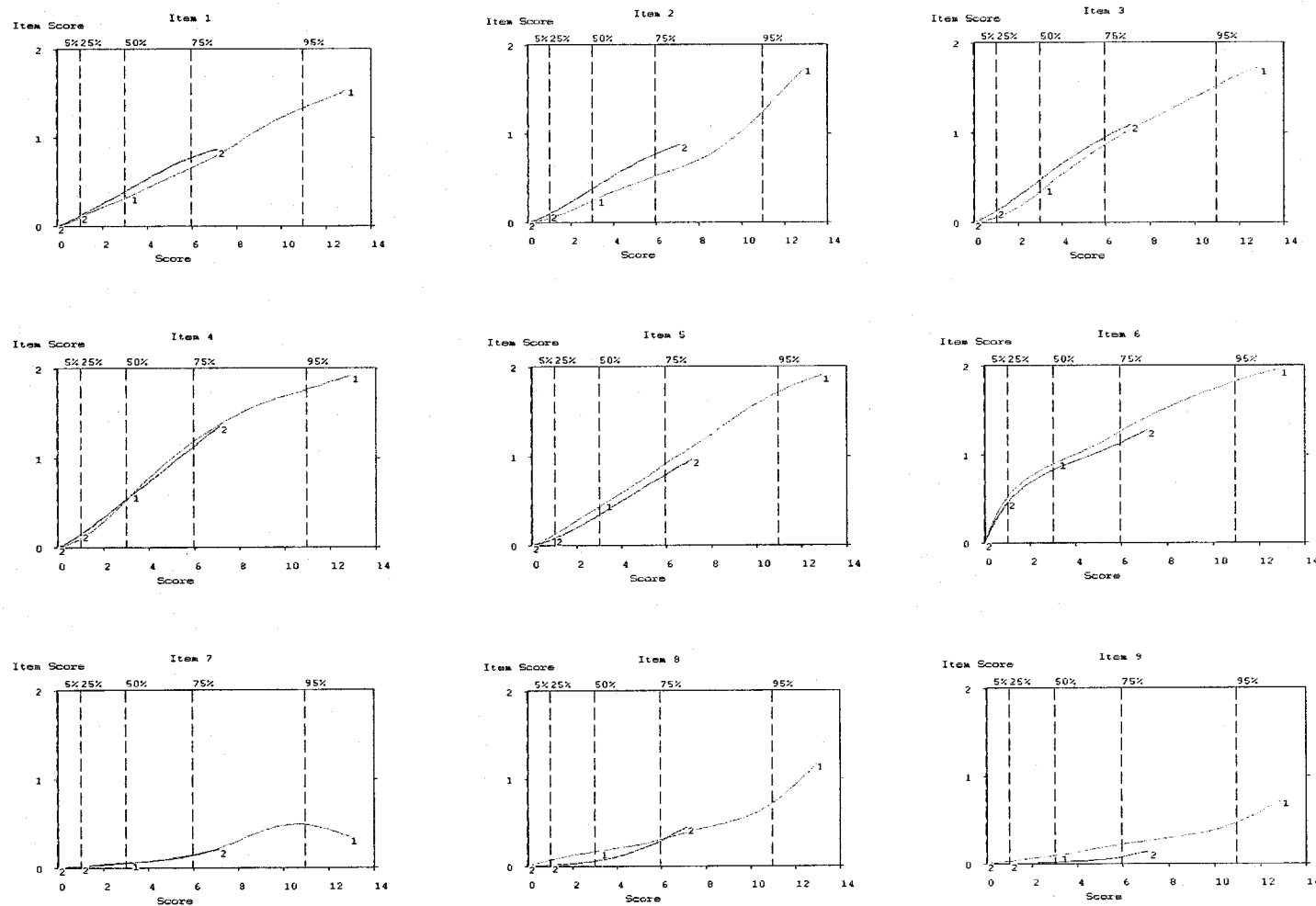


Figure 18.

Item characteristic curves for each item on the Managing Mood scale on the BCFPI among the First Nations clinical sample and the non-First Nations community sample.